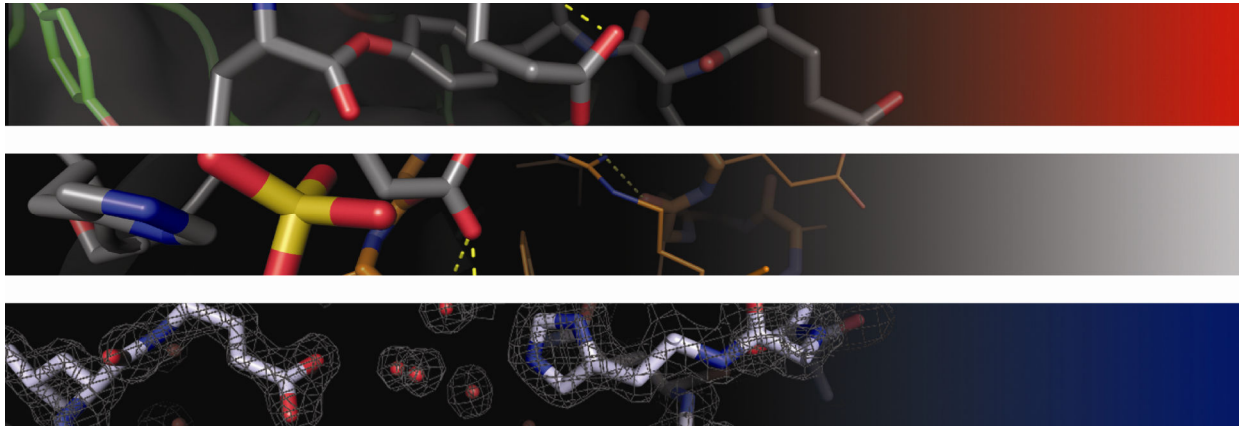


# Arbeiten zur strukturellen Charakterisierung von Thioesterase-, Kondensations- und Epimerisierungsdomänen nichtribosomaler Peptidsynthetasen



Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

dem Fachbereich Chemie der Philipps-Universität Marburg vorgelegt von

**Stefan-Andreas Samel** aus Ludwigshafen am Rhein

Marburg/Lahn, Juli 2009



# **Arbeiten zur strukturellen Charakterisierung von Thioesterase-, Kondensations- und Epimerisierungsdomänen nichtribosomaler Peptidsynthetasen**

Dissertation zur Erlangung des  
Doktorgrades der Naturwissenschaften  
(Dr. rer. nat.)

dem  
Fachbereich Chemie  
der Philipps-Universität Marburg

vorgelegt von

**Stefan-Andreas Samel**

aus Ludwigshafen am Rhein

Marburg/Lahn, Juli 2009

Vom Fachbereich Chemie der Philipps-Universität Marburg  
als Dissertation am 10. Juli 2009 angenommen.

Erstgutachter:	Herr Prof. Dr. L.-O. Essen
Zweitgutachter:	Herr Prof. Dr. M. A. Marahiel
Tag der Disputation:	31. August 2009



*meiner Familie*



Ein Teil der vorliegenden Arbeit wurde bereits in den folgenden Publikationen veröffentlicht:

- **Stefan A. Samel, Björn Wagner, Mohamed A. Marahiel, Lars-O. Essen** (2006)  
*The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-ribosomal Lipopeptide.* J. Mol. Biol. **359**, 876-89.
- **Stefan A. Samel, Georg Schoenafinger, Thomas A. Knappe, Mohamed A. Marahiel, Lars-O. Essen** (2007) *Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase.* Structure **15** (7), 781-92.

Weitere Publikationen:

- **Stefan A. Samel, Mohamed A. Marahiel, Lars-O. Essen** (2008) *How to tailor non-ribosomal peptide products – new clues about the structures and mechanisms of modifying enzymes.* Mol. BioSyst. **4**, 387-393.
- **Alan Tanovic, Stefan A. Samel, Lars-O. Essen, Mohamed A. Marahiel** (2008) *Crystal Structure of the Termination Module of a Nonribosomal Peptide Synthetase.* Science **321**, 659-663.
- **Verena Helmetag, Stefan A. Samel, Michael G. Thomas, Mohamed A. Marahiel, Lars-O. Essen** (2009) *Structural basis for the erythro-stereospecificity of the L-arginine oxygenase VioC in viomycin biosynthesis.* FEBS J. (in print)



## Zusammenfassung

Nichtribosomal synthetisierte Peptide stellen eine breite Klasse von Naturstoffen dar und besitzen meist pharmakologisch nützliche Eigenschaften. Diese von Bakterien und Pilzen produzierten, niedermolekularen Peptidverbindungen enthalten meist modifizierte Reste, zum Teil Aryl- und Fettsäuren und sind häufig makrozyklisiert. Die strukturelle Vielfalt wird durch die Ribosomen-unabhängige Synthese an großen Enzymen bzw. Enzymkomplexen, den Nichtribosomalen Peptid-synthetasen (NRPS), möglich. Sie sind ähnlich den Typ I Fettsäuresynthetasen modular aufgebaut, wobei definierte Domänen die erforderlichen Reaktionen katalysieren.

Die vorliegende Arbeit beschäftigt sich mit der strukturellen Aufklärung der Domänenstrukturen von Thioesterase-, Kondensations- und Epimerisierungsdomänen. Thioesterasedomänen liegen meist am Ende eines Syntheseclusters und katalysieren die Produktabspaltung. Dies geschieht häufig unter Makrozyklisierung des Produktpeptids. Kondensationsdomänen hingegen katalysieren die Ausbildung von Peptidbindungen in den Peptidprodukten. Epimerisierungsdomänen epimerisieren Aminosäurereste in den enzymgebundenen Intermediaten. In Synthetaseclustern sind sie darüber hinaus an den intramolekularen Wechselwirkungen zur gegenseitigen Erkennung aufeinanderfolgender Synthetasen beteiligt.

Für die Thioesterase des Fengycin Syntheseclusters aus *B. subtilis* wurde durch Lösung ihrer Struktur bei 1.8 Å Auflösung und mit Hilfe computergestützter Methoden ein Strukturmodell eines Enzym-Produkt-Komplexes entwickelt. In diesem Komplex ist das Produkt *edge-on* in einer grabenförmigen Vertiefung der Proteinoberfläche gebunden. Die Stabilität des Komplexes konnte mit Hilfe einer Moleküldynamik-Simulation über einen Zeitraum von 10 ns gezeigt werden. Anschließend durchgeführte Enzymaktivitätstests unter Verwendung gezielt veränderter Substrat-peptide unterstützen das Modell des Enzym-Produkt-Komplexes.

Die 1.8 Å Struktur eines PCP-C Bidomänenproteins aus der Tyrocidin Synthetase TycC aus *B. brevis* wurde mittels multipler, anomaler Dispersion (MAD) gelöst. Sie liefert Einblick sowohl in die Struktur einer internen Kondensationsdomäne als auch in die Lage des Domänen-verbindenden Linkers. Die beobachtete Konformation der *Peptidyl Carrier Protein*-Domäne entspricht der A/H-Konformation. Die Kondensationsdomäne besteht aus zwei Subdomänen, die jeweils Ähnlichkeit zum Chloramphenicol-Acetyltransferase (CAT)-Faltungstyp aufweisen. An der Grenzfläche der Subdomänen liegen sowohl das katalytische Zentrum als auch zwei Bereiche, über die die beiden Subdomänen miteinander interagieren. Dies sind zum einen die *bridge region*, die das aktive Zentrum überbrückt, zum anderen der *floor loop*, der den Boden des aktiven Zentrums bildet und dessen Wechselwirkungen zur strukturellen Integrität des aktiven Zentrums beitragen. Der Mechanismus der Kondensationsreaktion ist noch nicht vollständig geklärt. Strukturbasierte Berechnungen deuten darauf hin, dass, entgegen der bisherigen Modellvorstellung, der Histidinrest

im aktiven Zentrum nicht als katalytische Base agiert. Auf Basis weiterer in der Kristallstruktur beobachteter Wechselwirkungen wird ein neues Modell für die Katalyse der Peptidbindungs-bildung in Kondensationsdomänen vorgeschlagen.

Die strukturelle Charakterisierung der NRPS-Epimerisierungsdomäne der Tyrocidin Synthetase TycA aus *B. brevis* bei 1.65 Å Auflösung gibt erstmals Einblicke in die räumliche Organisation dieser Kofaktor-unabhängigen Epimerasen. Bedingt durch die Verwandtschaft mit den Kondensationsdomänen bestehen die Epimerisierungsdomänen ebenfalls aus zwei CAT-ähnlichen Subdomänen. Ein Vergleich mit der oben beschriebenen Kondensationsdomäne zeigt auch deutliche Unterschiede im Bereich der *bridge region*, des *floor loops* sowie im aktiven Zentrum.

Die Epimerisierungsdomäne enthält in ihrem aktiven Zentrum neben einem konservierten, katalytisch aktiven Histidinrest, H146, ebenfalls einen konservierten Glutamatrest, E285. Auch für die Epimerisierungsdomäne ergeben die Berechnungen der Protonierungszustände, dass H146 protoniert vorliegt. Dies widerspricht den bisherigen Modellvorstellungen, wonach H146 als katalytische Base auftritt. Während der Mechanismus ungeklärt bleibt, wird die Epimerisierungsdomäne mit strukturell und funktionell ähnlichen Proteinen verglichen, um Hinweise auf den Katalysemechanismus zu erhalten.

## Summary

Most bacteria and fungi produce secondary metabolites, including nonribosomally synthesized peptides, some of which possess potent pharmacological activities. These peptidic compounds comprise 3 to 15 'amino acids' and often contain unusual building blocks such as aryl or fatty acids or non-proteinogenic amino acids, respectively. These modifications strongly add to their structural as well as functional diversity and are made possible by means of large enzymes or enzyme complexes, the so-called nonribosomal peptide synthetases (NRPS). NRPS exhibit a modular organization with each module containing specific domains catalyzing the basic reactions required for peptide assembly.

The goal of this work is the elucidation of protein structures of catalytic domains involved in nonribosomal peptide synthesis with the main focus being set on thioesterase domains, condensation domains as well as epimerization domains. Thioesterase domains catalyze the release of the enzyme-bound, full-length peptide and mostly the concomitant cyclization of the product. Condensation domains catalyze peptide bond formation between two aminoacyl intermediates or an aminoacyl and a peptidyl intermediate, respectively. Epimerization domains primarily alter the stereochemical configuration of amino acid residues in the enzyme-bound intermediates. In bacterial synthetase clusters, they are also involved in intermolecular interactions leading to mutual recognition of cognate synthetases.

After the elucidation of the 1.8 Å crystal structure of the thioesterase domain involved in fengycin biosynthesis, a model of a putative enzyme-substrate complex was derived *in silico* using a combination of docking and molecular dynamics. The model suggests an *edge-on* binding mode of the peptide in the enzyme's binding pocket. Finally the binding mode was tested biochemically through analysis of the enzyme's acceptance towards altered substrates as well as by molecular dynamics simulations for stability of the complex.

The 1.8 Å crystal structure of the PCP-C bidomain protein of the tyrocidin synthetase TycC from *B. brevis* was solved using multiple anomalous dispersion (MAD). This structure provides insight not only into the architecture of an internal condensation domain but also into the orientation and interactions of the linker connecting the two adjacent domains. The conformation of the PCP domain corresponds to the A/H-conformation. The condensation domain comprises two chloramphenicol acetyl transferase (CAT)-like subdomains with its active site being located at the two subdomains' interface. Interactions between these two subdomains are mainly observed in two regions, the *bridge region* and the *floor loop*. The former bridges the active site, the latter forms the bottom of the domain's active site and contributes to the active site's structural integrity through an intricate network of hydrogen bonds and salt bridges. The exact mechanism of the condensation reaction remains unknown. However, structure-based calculations contradict the model in which

the active site histidine acts as a catalytic base. Lead by further analysis of the active site's surroundings a new model for the catalytic mechanism is described.

The structure of an NRPS-epimerization domain from the tyrocidin synthetase TycA from *B. brevis* was solved at 1.65 Å resolution and gives a first impression of the structural organization of these cofactor-independent epimerases. Due to their evolutionary relatedness to condensation domains they also consist of two CAT-like subdomains. Direct comparison of condensation and epimerization domains indicates profound differences in their active sites, their *bridge regions* and their *floor loops*. In addition to the conserved, catalytically active histidine residue, H146, the active site of the epimerization domain contains a likewise conserved glutamate residue, E285. As for the condensation domain, calculations regarding the protonation states result in a protonated H146, thus precluding a role as a catalytic base as described in earlier models. While the detailed mechanism remains unknown, the epimerization domain is compared to structurally and functionally similar enzymes in order to improve our understanding of the epimerization mechanism.







# Inhaltsverzeichnis

<b>1. Einleitung .....</b>	<b>1</b>
1.1 Die nichtribosomale Peptidsynthese und ihre Produkte .....	1
1.2 Mechanismus und Logik der nichtribosomalen Peptidsynthese .....	4
1.3 Essentielle Domänen .....	7
Adenylierungsdomäne (A) .....	7
Peptidyl Carrier Domäne (PCP) .....	9
Kondensationsdomäne (C) .....	12
1.4 Optionale und modifizierende Domänen .....	15
1.5 Fettsäuresynthasen und Polyketidsynthasen .....	24
1.6 Aufgabenstellung der vorliegenden Arbeit .....	26
 <b>2. Material .....</b>	 <b>27</b>
2.1 Geräte .....	27
2.2 Chemikalien, Biochemikalien und Enzyme .....	29
2.2.1 Puffer .....	29
2.3 Medien .....	30
2.3.1 Luria-Bertani-Medium (LB-Medium) .....	30
2.3.2 LB-Agar-Festmedium .....	30
2.3.3 M9-Minimalmedium .....	31
2.4 Vektoren .....	32
pQE-60 .....	32
pQE-61 .....	32
2.5 Bakterienstämme .....	33
2.5.1 Klonierungsstämme .....	33
2.5.2 Expressionsstämme .....	33
 <b>3 Methoden .....</b>	 <b>35</b>
3.1 Molekularbiologische Methoden .....	35
3.1.1 Synthese von DNA-Oligonukleotiden .....	35
3.1.2 Polymerase-Kettenreaktion (PCR) .....	35
3.1.3 Inverse Polymerase-Kettenreaktion (iPCR) .....	35
3.1.4 Reinigung von PCR-Fragmenten .....	36
3.1.5 Mini-Plasmidpräparation mit QIAprep® spin-Säulen .....	36
3.1.6 Hydrolytische Spaltung von DNA .....	37
3.1.7 Ligation .....	37

3.1.8 Entsalzung von Ligationsansätzen .....	37
3.1.9 Agarosegelelektrophorese .....	37
3.1.10 Photometrische DNA-Konzentrations- und Reinheitsbestimmung .....	38
3.1.11 DNA-Sequenzanalyse .....	38
3.1.12 Darstellung und Transformation kompetenter <i>E. coli</i> -Zellen .....	38
3.1.13 Stammhaltung .....	39
<b>3.2 Proteinchemische Methoden .....</b>	<b>39</b>
3.2.1 Heterologe Genexpression in <i>E. coli</i> .....	39
3.2.2 Präparation von <i>E. coli</i> Rohzelextrakten .....	40
3.2.3 Affinitätschromatographie .....	40
3.2.4 Ultrafiltration .....	40
3.2.5 Gelfiltration .....	41
3.2.6 Entsalzung und Umpufferung .....	41
3.2.7 Sterilfiltration .....	41
3.2.8 Proteinkonzentrationsbestimmung .....	42
3.2.9 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE) .....	42
<b>3.3 Proteinkristallisation und Strukturbestimmung .....</b>	<b>42</b>
3.3.1 <i>hanging drop</i> - und <i>sitting drop</i> -Kristallisation .....	43
3.3.2 <i>Soaking</i> -Experimente .....	43
3.3.3 <i>Seeding</i> -Experimente .....	44
3.3.4 Testen von Cryo-Bedingungen .....	44
3.3.5 Datensammlung und Datenreduktion .....	44
3.3.6 Suche nach Sequenz- und Strukturhomologen .....	45
3.3.7 Strukturlösung mittels <i>molecular replacement</i> .....	45
3.3.8 Strukturlösung mittels multipler, anomaler Dispersion (MAD) .....	45
3.3.9 Modellbau und Strukturverfeinerung .....	46
3.3.10 Moleküldynamische Simulation .....	46
3.3.11 Berechnung von Protonierungszuständen .....	46
3.3.12 Graphische Darstellungen .....	47
<b>4 Ergebnisse .....</b>	<b>49</b>
<b>4.1 Struktur-Funktions-Beziehungen einer NRPS-Thioesterasedomäne .....</b>	<b>50</b>
4.1.1 Verfeinerung der Struktur der PMSF-inhibierten Fengyc in Thioesterasedomäne .....	51
4.1.2 <i>Soaking</i> mit substratanalogen Verbindungen .....	53
4.1.3 <i>Modelling</i> und Moleküldynamik Simulation eines FenTE-Fengycin-Komplexes .....	55
4.1.4 <i>Docking</i> und moleküldynamische Simulation des Enzym-Produkt Komplexes .....	56
<b>4.2 Kristallisation und Strukturlösung einer PCP-C Bidomäne .....</b>	<b>58</b>
4.2.1 Expression und Aufreinigung der nativen PCP-C Bidomäne .....	59
4.2.2 Kristallisation der nativen PCP-C Bidomäne .....	59
4.2.3 Datensammlung von nativem PCP-C .....	60
4.2.4 Kristallisation von Selenomethionin-markiertem PCP-C .....	61

4.2.5 MAD-Datensammlung von Selenomethionin-markiertem PCP-C .....	61
4.2.6 Strukturlösung der Bidomäne PCP-C .....	61
4.3 Arbeiten zur Strukturaufklärung von Epimerisierungsdomänen .....	64
4.3.1 Die C-terminal verkürzte Epimerisierungsdomäne der Tyrocidin Synthetase A .....	65
Klonierung und Expression von <i>tycA-E</i> ( <i>kurz</i> ) .....	65
Kristallisation der verkürzten Epimerisierungsdomäne der Tyrocidin Synthetase A .....	67
Produktion der Selenomethionin-markierten Epimerisierungsdomäne aus TycA .....	68
Kristallisation der Selenomethionin-markierten Epimerisierungsdomäne aus TycA .....	68
Datensammlung der Epimerisierungsdomäne aus Tyrocidin Synthetase A .....	68
Strukturlösung der TycA-Epimerisierungsdomäne .....	70
4.3.2 Die Epimerisierungsdomäne der Tyrocidin Synthetase B .....	71
Stand der Forschung zu Beginn des Projekts .....	71
Optimierung der Kristallisationsbedingung und Datensammlung .....	71
Aufnahme von Datensätzen .....	72
4.3.3 Weitere bearbeitete Epimerisierungsdomänen .....	74
<b>5 Diskussion .....</b>	<b>75</b>
5.1 Das <i>edge-on</i> Modell des Fengycin Thioesterase-Fengycin-Komplexes .....	75
Biochemische Überprüfung des Komplexmodells .....	79
5.2 Faktoren der Substratbindung in NRPS-Thioesterasedomänen .....	80
Der Einfluss des intramolekularen Nukleophils .....	82
Wechselwirkungen zwischen PCP- und TE-Domänen .....	82
5.3 Die Struktur der PCP-C-Bidomäne .....	84
5.4 Wechselwirkungen zwischen der C-Domäne, ihrer donorseitigen PCP-Domäne und dem Domänen-verbindenden Linker-Bereich .....	89
5.5 Das aktive Zentrum der Kondensationsdomäne .....	92
Die <i>bridge region</i> und der <i>floor loop</i> .....	94
Vergleich der ribosomalen und nichtribosomalen Peptidbindungsbildung .....	96
5.6 Die Struktur der TycA-Epimerisierungsdomäne .....	98
5.7 Das aktive Zentrum der TycA-Epimerisierungsdomäne .....	102
Strukturelle Analyse ausgewählter Mutationen in NRPS-Epimerisierungsdomänen .....	103
Mechanistische Aspekte der Epimerisierungsreaktion .....	107
5.8 Vergleich der Kondensations- und Epimerisierungsdomänen in der nichtribosomalen Peptidsynthese .....	111
Mechanistische Ähnlichkeiten zwischen Kondensations- und Epimerisierungsreaktion .....	112
<b>6 Ausblick .....</b>	<b>115</b>

<b>7 Literaturverzeichnis .....</b>	<b>117</b>
 <b>8 Anhang .....</b>	 <b>125</b>
8.1 Verwendete Abkürzungen .....	125
8.2 Abkürzungen proteinogener und nichtproteinogener Aminosäuren .....	128
8.3 <i>Core</i> -Motive der katalytischen Domänen nichtribosomaler Peptidsynthetasen .....	129
8.4 Verwendete Kristallisationslösungen .....	130
8.5 Kristallisation von NRPS-Epimerisierungsdomänen .....	135
8.5.1 Kristallisation C-terminal verkürzter NRPS-Epimerisierungsdomänen .....	135
8.6 Berechnete Matthews-Koeffizienten für die Kristallform von TycB3 E .....	137
8.7 Versuche der Strukturlösung via <i>molecular replacement</i> .....	138
8.8 Zugänge zum aktiven Zentrum der TycA-Epimerisierungsdomäne .....	142
 Danksagung .....	 144
 Erklärung .....	 146



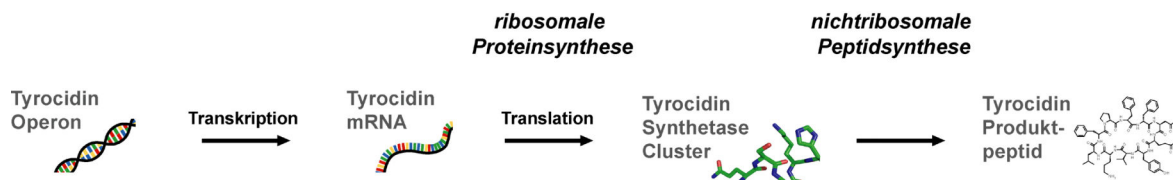




# 1 Einleitung

## 1.1 Die nichtribosomale Peptidsynthese und ihre Produkte

Eine Vielzahl von Bakterien und Pilzen produzieren bioaktive Polypeptid- und Polyketidverbindungen mit Hilfe nichtribosomaler Peptidsynthetasen (NRPS) bzw. Polyketidsynthetasen (PKS). Bei den nichtribosomalen Peptidsynthetasen handelt es sich um große multifunktionale Enzyme, die unabhängig von einer RNA-Matrize die Zusammensetzung der 3 bis 15 Aminosäuren großen Peptide katalysieren [Sieber, 2005]. Die Synthetasen und Synthasen selbst sind im Genom ihres jeweiligen Produzentenstammes kodiert. Diese Gene werden von einer RNA-Polymerase transkribiert und anschließend am Ribosom translatiert. Die nichtribosomale Peptidsynthese ist demnach, wie in Abbildung 1.1 dargestellt, eine Ergänzung zur ‚klassischen‘, ribosomalen Proteinsynthese.



**Abbildung 1.1: Rolle der nichtribosomalen Peptidsynthese neben der klassischen, ribosomalen Proteinsynthese**

Die Synthetasen sind im Genom des Produzenten kodiert und werden transkribiert und am Ribosom translatiert. Die katalytische Funktion der Synthetasen besteht in der Produktion eines Peptids – in diesem Beispiel des zyklischen Peptid-antibiotikums Tyrocidin.

Die nichtribosomal synthetisierten Peptide bzw. Polyketide sind als Sekundärmetabolite für ihren jeweiligen Produzenten nicht essentiell. Sie verfügen aber über Eigenschaften, die ihrem Produzenten einen Selektionsvorteil verleihen können. Mit der Entdeckung, dass diese funktionell und strukturell vielseitigen Peptid- und Polyketidverbindungen auch pharmakologisch interessante Eigenschaften besitzen, rückten sie in das Blickfeld der pharmazeutisch-medizinischen Forschung [Butler, 2006]. Die Wirkungen der beiden Naturstoffklassen erstrecken sich von antibiotisch über antiviral und immunosuppressiv bis zu zytostatisch. Heute werden diese Substanzen in der pharmazeutisch-medizinischen Forschung entweder direkt als Wirkstoffe eingesetzt oder sie dienen als Leitstruktur für Neu- und Weiterentwicklungen.

Die Notwendigkeit zur ständigen Entdeckung neuer und der Weiterentwicklung bekannter Naturstoffverbindungen lässt sich am Beispiel der Antibiotika anschaulich darstellen. Seit der Entdeckung des Penicillins und der anschließenden Erforschung weiterer Antibiotika sind die Risiken schwerer bakterieller Infektionen stetig reduziert worden. Allerdings treten in der Folge der Einführung eines Antibiotikums in die klinische Praxis unweigerlich resistente Erregerstämme auf [Walsh, 2000]. Zur Bekämpfung dieser Bakterienstämme sind andere, noch wirksame Antibiotika

## 1 Einleitung

erforderlich. Die ständige Erforschung und Entwicklung neuer, wirksamer Antibiotika dient somit der Erhaltung unserer eigenen Gesundheit durch die Sicherung der Standards in der modernen medizinischen Versorgung.

Eine Auswahl nichtribosomal produzierter Peptide ist in Abbildung 1.2 dargestellt. Die Verbindungen sowie ihre Wirkungen und Einsatzgebiete sollen nun vorgestellt werden.

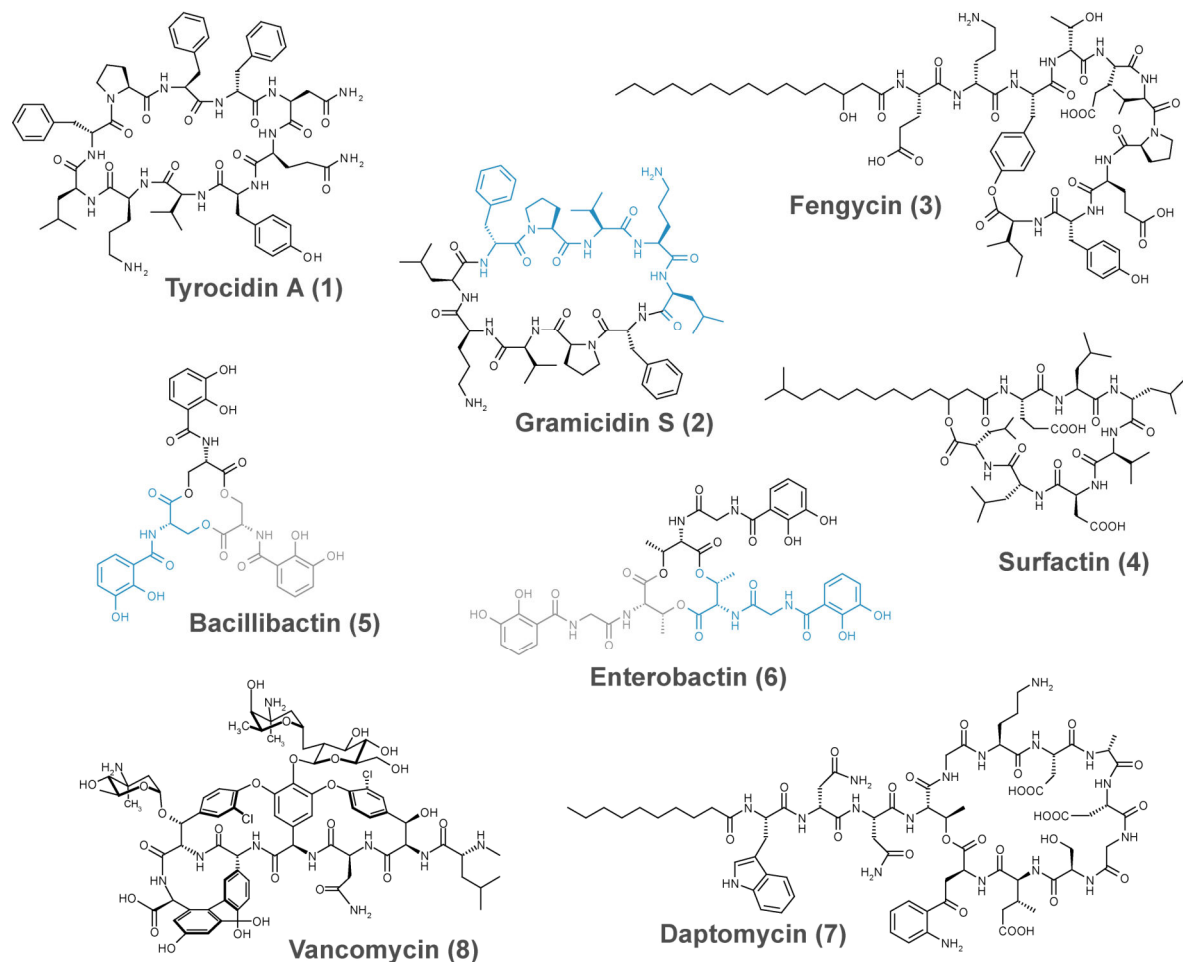
Neben dem Penicillin, welches selber ein Produkt der nichtribosomalen Peptidsynthese ist, sind das Tyrocidin A [Mootz, 1997] und das Gramacidin S [Kratzschmar, 1989] aus *Bacillus brevis* zwei Vertreter der Peptidantibiotika. Bei beiden Verbindungen handelt es sich um zyklische Dekapeptide, wobei sich das Gramacidin aus zwei identischen Pentapeptideinheiten zusammensetzt, was in Abbildung 1.2 durch die unterschiedliche Schattierung der zwei Hälften hervorgehoben ist. Die Konformation wird in beiden Fällen durch die Ausbildung eines intramolekularen, antiparallelen  $\beta$ -Faltblattes stabilisiert. Ebenso enthalten beide Antibiotika die nichtproteinogene Aminosäure Ornithin sowie D-konfigurierte Phenylalaninreste. Sie werden gegen Infektionen durch gram-positive Bakterien eingesetzt, wobei sie deren Membranintegrität stören und dadurch Ionengradienten zerstören.

Surfactin und Fengycin sind verzweigt-zyklische Lipopeptide aus *Bacillus subtilis*. Surfactin ist ein Biodetergens, verfügt darüber hinaus aber auch über antibakterielle und antivirale Eigenschaften [Peypoux, 1999]. Das Fengycin-analoge Plipastatin inhibiert die Phospholipase A<sub>2</sub> [Umezawa, 1986] und wirkt hauptsächlich gegenüber filamentösen Pilzen antibiotisch [Vanittanakom, 1986]. Beide Lipopeptide werden von ihren Produzenten sekretiert und können durch ihre unpolaren Fettsäureanteile mit den Membranen anderer Zellen interagieren. Surfactin kann entweder Poren ausbilden, welche essentielle Kationengradienten ausgleichen, oder aufgrund seiner Detergenseigenschaften die Membranen solubilisieren, was ebenfalls zum Zelltod führt.

Bei Bacillibactin aus *Bacillus subtilis* [May, 2002] und Enterobactin aus *Escherichia coli* [Gehring, 1998] handelt es sich um Trimere, die sich aus 2,3-Dihydroxybenzoyl-Seryl- bzw. 2,3-Dihydroxybenzoyl-Glycyl-Threonyl-Einheiten zusammensetzen. Sie gehören zur Gruppe der Siderophore, die mit ihren sechs catecholischen, oktaedrisch angeordneten Hydroxidgruppen Fe<sup>3+</sup>-Ionen sehr stark komplexieren können [Crosa, 2002; Miethke, 2007]. Sie werden bei Eisenmangel synthetisiert und aus dem produzierenden Organismus exportiert. In der Umgebung oder im Wirt vorhandene Eisenionen werden gebunden und nach Aufnahme des Komplexes in die Zelle und nach Hydrolyse des Siderophors dem Produzenten verfügbar gemacht. Auf diese Weise stellt der Produzent seinen Bedarf an Eisenionen sicher und erhöht seine Überlebenschancen in Zeiten des Mangels.

Das Daptomycin aus *Streptomyces roseosporus* [Kirkpatrick, 2003] gehört zur Klasse der aziden Lipopeptiden, der auch die Antibiotika CDA (calcium-dependent antibiotic) aus *Streptomyces coelicolor* [Hojati, 2002] und A54145 aus *Streptomyces fradiae* [Fukuda, 1990] angehören. Auch diese Gruppe entfaltet ihre antibiotische Wirkung gegenüber gram-positiven Bakterien, indem sie sich in deren Membranen einlagern. Es wird vermutet, dass die  $\text{Ca}^{2+}$ -Ionen die Interaktion mit der Membran vermitteln. Daptomycin wurde im Jahr 2003 als erster Vertreter dieser strukturell neuen Antibiotika-Klasse von der *Food and Drug Administration* für den US-amerikanischen Markt zugelassen [Kirkpatrick, 2003].

Vancomycin aus *Amycolatopsis orientalis* ist ein Glykopeptid-Antibiotikum, das über mehrere Jahre als Reserveantibiotikum galt und zur Behandlung multiresistenter Erreger eingesetzt wurde (z. B. Oxacillin-resistente bzw. Methicillin-resistente *S. aureus*, ORSA/MRSA). Seit 1987 sind



**Abbildung 1.2: Auswahl nichtribosomal synthetisierter Peptide**

Strukturen der Produktpeptide Tyrocidin A (1) und Gramicidin S (2) aus *B. brevis*, Fengycin (3), Surfactin (4) und Bacillibactin (5) aus *B. subtilis*, Enterobactin (6) aus *E. coli*, Daptomycin (7) aus *S. roseosporus* und Vancomycin (8) aus *A. orientalis*. Die monomeren Einheiten des Dimers Gramicidin S sowie die der Trimere Bacillibactin und Enterobactin sind farblich unterschiedlich gekennzeichnet.

## 1 Einleitung

allerdings auch Vancomycin-resistente Erregerstämme bekannt, deren Auftreten insbesondere in Kliniken ein großes Problem darstellt [Murray, 2000]. Die Wirkung des Vancomycins gegenüber gram-positiven Bakterien basiert darauf, dass es in die Zellwandsynthese eingreift. Durch Bindung der Peptidoglykan-D-Alanyl-D-Alanin-Termini wird die nachfolgende Reaktion, die Transpeptidierung der Peptidoglykane mit den Pentaglycin-Einheiten, unterbunden, was zu einer instabilen bakteriellen Zellwand und schließlich zur osmotischen Lyse der Bakterienzelle führt [Walsh, 2000].

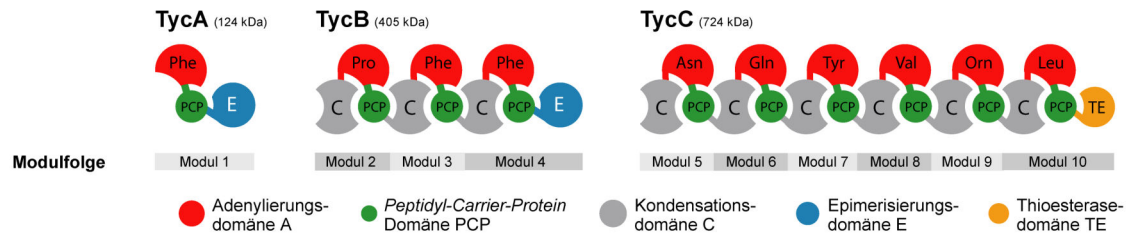
Aus den in Abbildung 1.2 dargestellten Verbindungen werden zudem zwei grundlegende Unterschiede gegenüber üblichen, ribosomal synthetisierten Proteinen deutlich. Einerseits finden sich in der nichtribosomalen Peptidsynthese häufig nichtproteinogene Aminosäuren. Diese hemmen den Abbau durch Proteasen und erlauben eine sehr genaue Anpassung der Peptidprodukte an die Moleküle, mit denen sie in Wechselwirkung treten müssen. Andererseits enthalten die meisten nichtribosomal synthetisierten Peptide makrozyklische Strukturen, wodurch die Anzahl der möglichen Konformationen im Produkt drastisch einschränkt ist. Die Bedeutung der Makrozyklisierung für die biologische Aktivität zeigt sich darin, dass lineare Analoga zyklischer Peptide weniger aktiv sind [Kohli, 2002].

Die nichtribosomalen Peptidsynthetasen sind in Hinblick auf ihre modulare Organisation und ihr Syntheseprinzip mit den Fettsäure- und Polyketidsynthetasen verwandt. Die Proteinstrukturen und die Synthesemodi dieser beiden Synthase-Klassen werden in Kapitel 1.5 vorgestellt.

## 1.2 Mechanismus und Logik der nichtribosomalen Peptidsynthese

Die Synthese der genannten funktionell und strukturell vielseitigen Verbindungen erfolgt an großen Enzymen bzw. Enzymkomplexen, den nichtribosomalen Peptidsynthetasen (NRPS). Während die Synthetasen in Pilzen meist aus einem einzigem Polypeptid bestehen, sind die katalytischen Einheiten der Synthetasen bakterieller Systeme oftmals auf mehrere Enzyme verteilt. Für die Synthese lagern sich die beteiligten Enzyme zu großen Enzymkomplexen (engl.: *cluster*) zusammen.

In Abbildung 1.3 ist dies am Beispiel der Tyrocidin Biosynthese verdeutlicht. Für die Synthese des Dekapeptids Tyrocidin kommen die drei Synthetasen TycA, TycB und TycC zusammen, um in einem fließbandähnlichen Prozess das Produkt aus den konstituierenden Aminosäuren zusammenzufügen. Die Synthese erfolgt ausschließlich in eine Richtung, die in Abbildung 1.3 einer Bewegung von links nach rechts entspricht. Folglich werden die Positionen der in Syntheserichtung



**Abbildung 1.3: Schematische Darstellung des Tyrocidin Synthetase Clusters aus *Bacillus brevis* ATCC 8185**

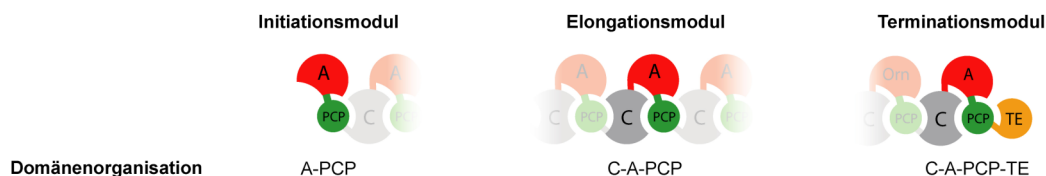
Das Cluster besteht aus drei Polypeptiden, den Synthetasen TycA, TycB und TycC. Das Synthetasecluster lässt sich in 10 Module unterteilen, die jeweils den Einbau einer Aminosäure in das Produktpeptid katalysieren. Die Module wiederum lassen sich weiter in Domänen unterteilen, die die Reaktionen katalysieren. Funktionelle und strukturelle Aspekte der einzelnen Domänentypen sind Gegenstand der Kapitel 1.3 und 1.4.

davorliegenden Einheiten als stromaufwärts (*upstream*) und nachfolgenden Einheiten als stromabwärts (*downstream*) benannt. Anhand ihrer Funktion lassen sich die Synthetasen in Module unterteilen, wobei jedes Modul den Einbau und ggf. die Modifikation einer Aminosäure katalysiert. Die Module lassen sich wiederum weiter in Domänen unterteilen, die die zugrundeliegenden chemischen Reaktionen katalysieren.

Zu Beginn der nichtribosomalen Peptidsynthese werden die Substrataminosäuren<sup>1</sup> von den Adenylierungsdomänen spezifisch erkannt und als Aminoacyladenylat aktiviert. Daraufhin wird die aktivierte Aminosäure auf die *Peptidyl Carrier Protein*-Domäne (PCP) übertragen. Diese verfügt über einen posttranslational angebrachten Kofaktor (4'-Phosphopantethein), an den die Substrate und Zwischenstufen während der Synthese als Thioester gebunden sind. In einer von der Kondensationsdomäne (C) katalysierten Reaktion werden die PCP-gebundenen Aminoacylintermediate über eine Peptidbindung miteinander verknüpft. In diesem Schritt wird das entstehende Peptidylintermediat auf die PCP-Domäne des folgenden Moduls übertragen. Mit den weiteren Elongationsschritten, die ebenfalls ausschließlich in Richtung der Kettenverlängerung verlaufen, wird das wachsende Peptidylintermediat jeweils auf die PCP-Domäne des nächsten Moduls transferiert. Am letzten Modul hat das Produktpeptid seine volle Länge erreicht und wird meist unter Zyklisierung von der Synthetase abgespalten. Der hier beschriebene Syntheseweg mit thioestergebundenen Substraten und wiederkehrenden Reaktionen in Elongationszyklen erinnert an die Fettsäuresynthese. Da bei der nichtribosomalen Peptidsynthese jedoch mehrere *Carrier*-Proteine benötigt werden, wurde dieser Synthesemodus als *multiple carrier thiotemplate*-Mechanismus bezeichnet [Stein, 1996].

<sup>1</sup> Die nichtribosomale Peptidsynthese ist bei weitem nicht auf die proteinogenen Aminosäuren beschränkt. Wie aus den in Abbildung 1.2 dargestellten Produkten ersichtlich, können wesentlich mehr Verbindungen als Bausteine verwendet werden. So gibt es Adenylierungsdomänen, die spezifisch D-Aminosäuren oder Arylsäuren aktivieren. Daneben verleihen die Fettsäuren in den Lipopeptiden diesen Verbindungen ihren amphiphilen Charakter. In der vorliegenden Arbeit sind diese ungewöhnlichen Bausteine inbegriffen, auch wenn im Text – der besseren Lesbarkeit wegen – nur Aminosäuren explizit erwähnt werden.

## 1 Einleitung



**Abbildung 1.4: Die drei Modultypen und ihre typischen Domänenfolgen**

Neben diesen Domänenorganisationen können noch weitere, modifizierende Domänen auftreten. Die im Terminationsmodul auftretende Thioesterase-Domäne (TE) bewirkt häufig die Abspaltung und Zyklisierung des Produkts.

Basierend auf ihrer Position im Biosynthesecoluster und in der Domänenorganisation unterscheidet man drei verschiedene Modultypen: das Initiationsmodul, das Elongationsmodul und das Terminationsmodul (Abbildung 1.4). Ein Initiationsmodul besteht mindestens aus einer A- und einer PCP-Domäne für die spezifische Aminosäureaktivierung und für die Bindung an das Enzym. Das Elongationsmodul enthält in seinem Grundgerüst neben der A- und der PCP-Domäne noch eine Kondensationsdomäne (Abbildung 1.4 Mitte). Diese Domäne ist für jeden Elongationsschritt erforderlich, da sie die Verknüpfung der Intermediate katalysiert. Im Initiationsmodul fehlt sie, da die erste Kondensationsreaktion erst zwischen den Aminoacylsubstraten der Module 1 und 2 auftritt, die von der C-Domäne des zweiten Moduls, eines Elongationsmoduls, katalysiert wird. Das Terminationsmodul stellt ein Elongationsmodul dar, welches eine zusätzliche Domäne zur Produktabspaltung enthält. Im Beispiel des Tyrocidin Synthetaseclusters wird diese Abspaltung von einer Thioesterase-Domäne unter gleichzeitiger Zyklisierung des Peptids katalysiert.

Die Funktionen der A-, PCP- und C-Domänen – die selektive Aktivierung, die Anbindung der Intermediate an die Synthetase und deren Transport sowie die Ausbildung von Peptidbindungen – sind so grundlegend für die nichtribosomale Peptidsynthese, dass diese drei Domärentypen als essentielle Domänen bezeichnet werden. Die anderen Domänen werden dagegen zur Gruppe der optionalen Domänen gezählt, die modifizierend auf die Intermediate einwirken. Dennoch ist auch der Beitrag der optionalen Domänen zur strukturellen sowie chemisch-funktionellen Vielfalt und damit auch für die Aktivität der Produkte jeweils essentiell.

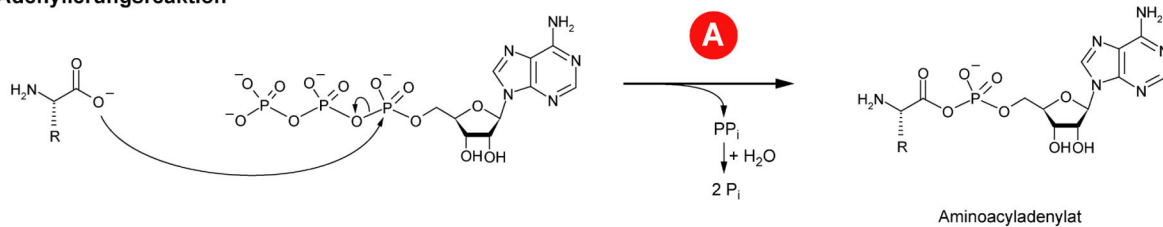
In den beiden folgenden Abschnitten werden zunächst die essentiellen Domänen und anschließend die optionalen Domänen vorgestellt. Da sich die vorliegende Arbeit vorrangig mit den Strukturen der Kondensations-, Thioesterase- und Epimerisierungsdomänen beschäftigt, fallen die Darstellungen zu diesen drei Domärentypen umfangreicher aus.

### 1.3 Essentielle Domänen

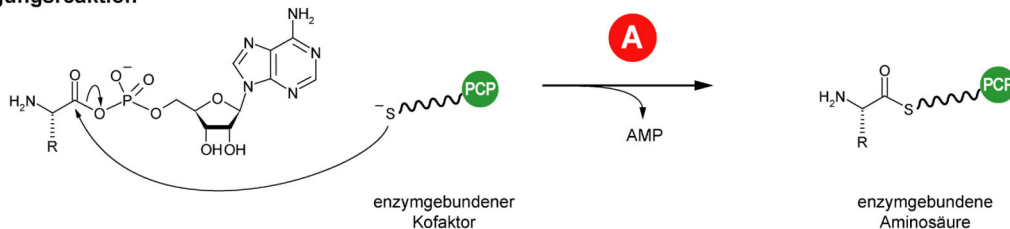
#### Adenylierungsdomäne (A-Domäne)

Adenylierungsdomänen aktivieren spezifisch die Aminosäure, die in das Produkt eingebaut werden soll. Die ATP-abhängige Aktivierung verläuft nach demselben Mechanismus wie die Reaktion der Aminoacyl-tRNA-Synthetasen und lässt sich ebenfalls in zwei Schritte unterteilen. Zunächst wird in der Adenylierungsreaktion ein gemischtes Anhydrid aus der Aminosäure und AMP gebildet. Im darauffolgenden Schritt wird die Aminosäure auf den enzymgebundenen Kofaktor 4'-Phosphopantethein (4'-Ppan) übertragen [Dieckmann, 1995]. Dabei ist die erste Reaktion prinzipiell reversibel. Erst durch die Hydrolyse des Nebenprodukts Pyrophosphat verläuft die Reaktion in Richtung des Aminoacyladenylats. Die beiden Teilreaktionen sind in Abbildung 1.5 gezeigt.

##### Adenylierungsreaktion



##### Übertragungsreaktion



#### Abbildung 1.5: Reaktionen der Adenylierungsdomäne zur Aminosäureaktivierung

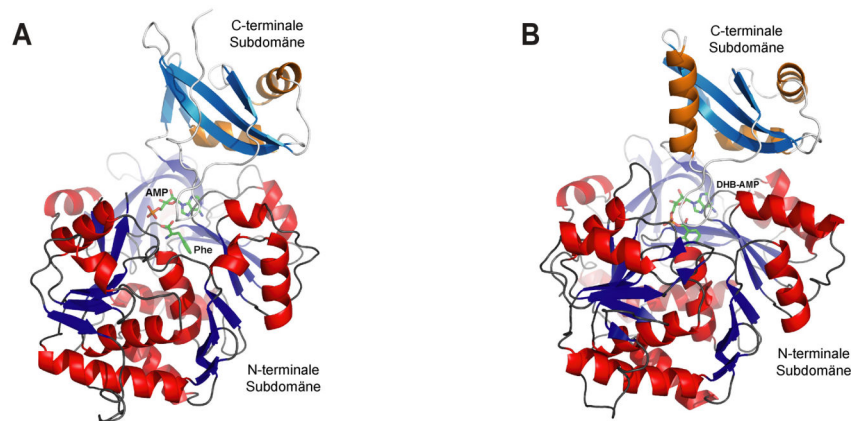
Die ausgewählte Aminosäure wird unter Verbrauch von ATP zu einem aktivierten Aminoacyladenylat umgesetzt. In einem zweiten Reaktionsschritt wird die Aminosäure als Thioester an den Kofaktor der PCP-Domäne gebunden. In der nichtribosomalen Peptidsynthese sind alle Substrate und Zwischenstufen über diese Kofaktoren gebunden.

Somit trägt die Spezifität der Adenylierungsdomänen maßgeblich zur Homogenität der Produktzusammensetzung bei. In diesem Punkt übernehmen A-Domänen und Aminoacyl-tRNA-Synthetasen vergleichbare Aufgaben, auch wenn die Spezifität der A-Domänen vergleichsweise gering ist. Trotz der Ähnlichkeit der katalysierten Reaktionen unterscheiden sich die beiden Enzymgruppen hinsichtlich ihrer Sequenz und Struktur. Sie stammen nicht von einem gemeinsamen Vorgänger ab. Vielmehr gehören die A-Domänen zu den ATP-verbrauchenden Enzymen des Luciferase-Typs [Conti, 1997].



## 1 Einleitung

Abbildung 1.6 zeigt die Kristallstrukturen zweier NRPS-Adenylierungsdomänen. Die A-Domäne der Synthetase GrsA, des Initiationsmoduls des Gramacidin S Synthetaseclusters aus *B. brevis* ATCC 9999, war die erste A-Domäne, deren Struktur bestimmt wurde. Das etwa 550 Aminosäuren umfassende Enzym besteht aus einer großen N-terminalen Subdomäne und einer kleineren C-terminalen Subdomäne. Das Substrat Phenylalanin und das Nebenprodukt des Kosubstrats, AMP, die beide ebenfalls in der Kristallstruktur beobachtet wurden, sind als Stabmodell dargestellt. In analoger Farbgebung ist in Abbildung 1.6B die A-Domäne der Synthetase DhbE dargestellt. Die größten strukturellen Unterschiede zwischen diesen beiden Strukturen liegen im Bereich der C-terminalen Subdomäne. So wird aus dem Vergleich der beiden abgebildeten Strukturen auch die Funktion der C-terminalen Subdomäne deutlich. Während die Struktur der GrsA-A-Domäne eine offene Konformation darstellt, wird in der Struktur der DhbE-A-Domäne eine Konformation beobachtet, in der das aktive Zentrum durch eine zusätzliche  $\alpha$ -Helix der C-terminalen Subdomäne verschlossen ist.



**Abbildung 1.6: Die Kristallstrukturen der Adenylierungsdomänen aus GrsA und DhbE**

(A) Die A-Domäne der Synthetase GrsA [Conti, 1997]. Die größere N-terminale Subdomäne (rot, blau und grau) bildet auch die Substratbindungstasche für das Substrat Phenylalanin und das Kosubstrat ATP aus. Die C-terminale Subdomäne ist in hellblau, orange und weiß dargestellt. (B) Die A-Domäne der Synthetase DhbE [May, 2002] mit dem gebundenen Zwischenprodukt, dem Adenylat der Dihydroxybenzoesäure. Die C-terminale Subdomäne der DhbE-A-Domäne bildet, verglichen mit der GrsA-A-Domäne, eine zusätzliche  $\alpha$ -Helix aus, die das aktive Zentrum verschließt.

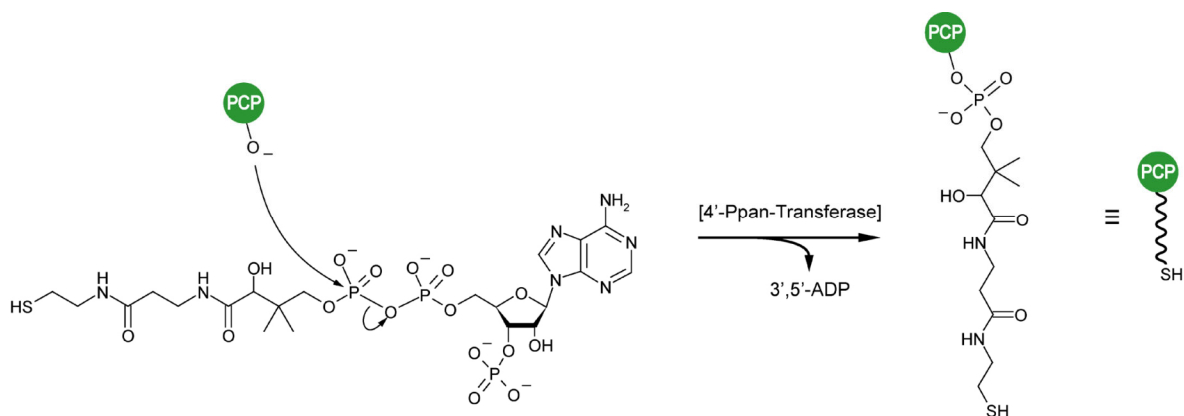
Die biochemische Charakterisierung der A-Domänen an gentechnisch isolierten A-Domänen ergab, dass die A-Domänen ihre ATP-abhängige enzymatische Aktivität und Substratspezifität beibehielten, auch wenn sie von den anderen Domänen ihrer Synthetase getrennt waren [Mootz, 1997]. Sequenzvergleiche ermöglichten die Identifizierung domänenspezifischer, konservierter Bereiche, der sogenannten *Core*-Motive. Die Sequenzen der *Core*-Motive der verschiedenen Domärentypen sind im Anhang wiedergegeben (Kapitel 8.3).



Die Kristallstruktur der A-Domäne von GrsA (Abbildung 1.6A) offenbarte ebenfalls, dass die Aminosäurereste der *Core-Motive* A4 und A5, die nahe am aktiven Zentrum liegen, die Spezifität der A-Domäne bestimmen. Das abgeleitete Modell, welches die Vorhersage der Aminosäurespezifität einer A-Domäne allein aufgrund ihrer Proteinsequenz erlaubt, wurde als ‚nichtribosomaler Code‘ bezeichnet [Stachelhaus, 1999] und durch gezielte Veränderung der Substratspezifität experimentell bestätigt [Eppelmann, 2002]. Nach Lösung der Kristallstruktur der Dihydroxybenzoat-aktivierenden A-Domäne aus der Synthetase DhbE des Bacillibactin-Clusters [May, 2002], ließ sich dieser Code auch auf Arylsäuren aktivierende A-Domänen ausweiten.

### Peptidyl Carrier Protein-Domäne (PCP-Domäne)

Die etwa 90 Aminosäuren umfassenden Domänen dienen der kovalenten Bindung der Aminoacyl- und Peptidylintermediate sowie deren Transport zwischen den katalytischen Domänen. Um diese Funktion erfüllen zu können, muss die PCP-Domäne durch posttranslationale Modifikation an einem konservierten Serinrest von ihrer katalytisch inaktiven *apo*-Form in die aktive *holo*-Form überführt worden sein. Dies geschieht durch die Übertragung eines 4'-Phosphopantethein-Kofaktors aus Coenzym A auf die Seitenkette des Serinrestes. Diese Reaktion wird von speziellen, im NRPS-Operon kodierten Phosphopantetheinyl-Transferasen (4'-Ppan-Transferasen) katalysiert [Lambalot, 1996].



**Abbildung 1.7: Posttranslationale Modifikation der PCP-Domäne mit dem Kofaktor 4'-Phosphopantethein**

Die Übertragung des 4'-Phosphopantethein aus Coenzym A auf den konservierten Serinrest der PCP-Domäne wird durch 4'-Phosphopantethein-Transferasen katalysiert. Der Kofaktor wird auch häufig durch eine gewellte Linie symbolisiert.

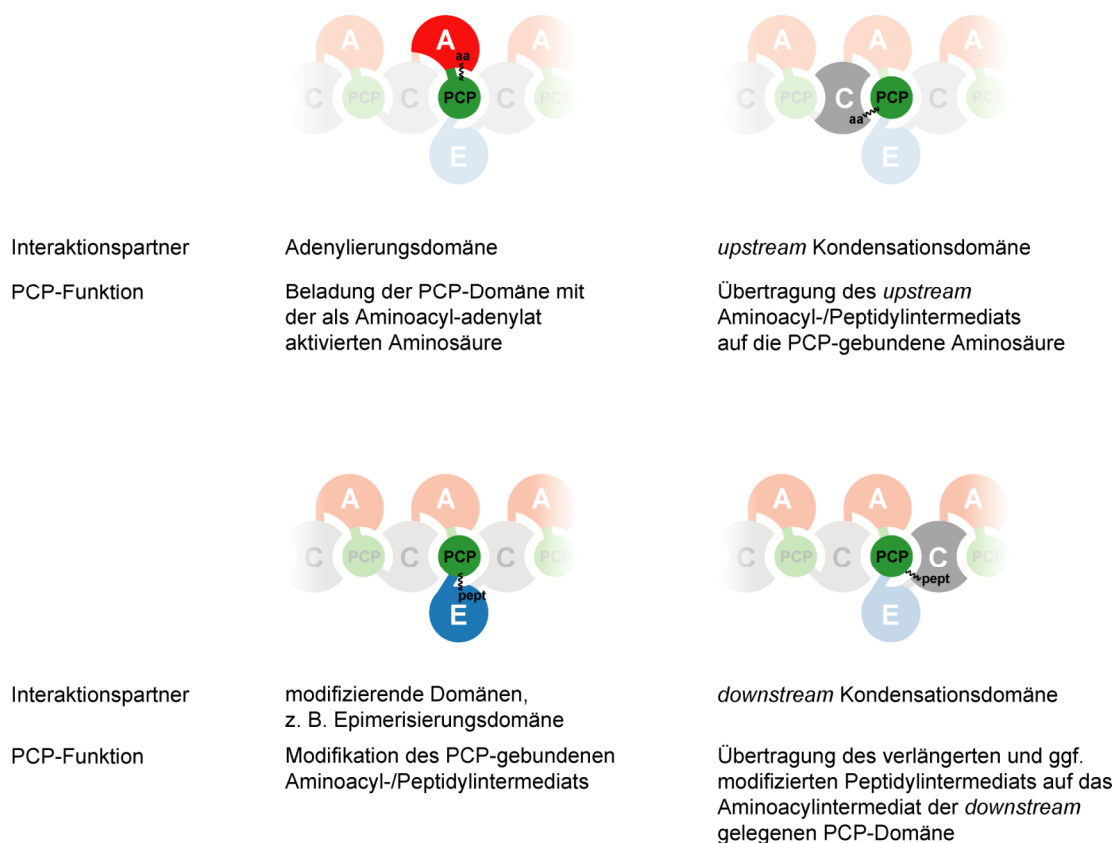
Es stellte sich heraus, dass die Phosphopantethein-Transferase Sfp des Surfactin Biosyntheseklusters neben dem eigentlichen Substrat CoA auch acylierte CoA-Derivate akzeptiert und auf den Serinrest überträgt. Damit war es möglich, synthetisierte Aminoacyl- [Belshaw, 1999] bzw. Peptidylintermediate [Sieber, 2003] auf die PCP-Domänen zu laden. Losgelöst von der Substrat-

## 1 Einleitung

spezifität der Adenylierungsdomänen und den davorliegenden Modulen ließ sich durch Fehlbeladung das Ausmaß der Substratspezifität für verschiedene Domärentypen bestimmen.

Wie bei der Behandlung der Adenylierungsdomänen beschrieben, wird die PCP-Domäne im zweiten Teilschritt der von der Adenylierungsdomäne katalysierten Reaktion mit einer Aminosäure beladen. Diese ist als Thioester an die Sulfhydrylgruppe des Kofaktors gebunden.

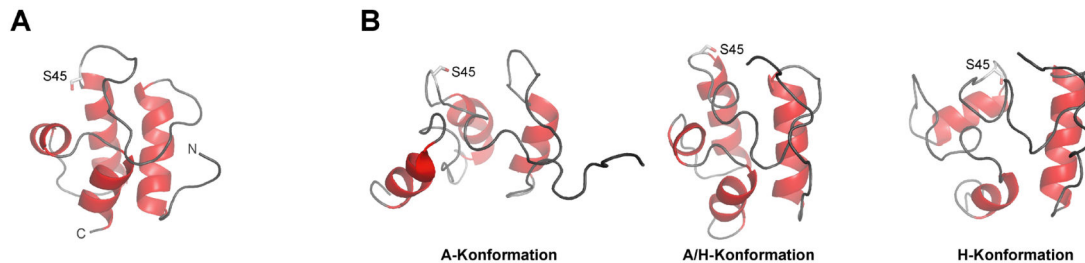
In einem Elongationsmodul tritt die PCP-Domäne danach zunächst mit der C-Domäne des davorliegenden Moduls in Wechselwirkung, wo das *upstream* synthetisierte Intermediat auf die PCP-gebundene Aminosäure übertragen wird. Das verlängerte, enzymgebundene Peptidylintermediat kann durch modifizierende Domänen oder durch *in trans* agierende Enzyme verändert werden. Abschließend tritt die PCP-Domäne mit der *downstream* gelegenen C-Domäne in Wechselwirkung, wo das Peptidylintermediat auf die enzymgebundene Aminosäure der nachfolgenden, benachbarten PCP-Domäne weitergereicht wird. Die zentrale Rolle der PCP-Domäne im Syntheseprozess mit den einzelnen Interaktionspartnern in den verschiedenen Reaktionsschritten ist in Abbildung 1.8 dargestellt.



**Abbildung 1.8: Typische Interaktionen einer Peptidyl Carrier Protein-Domäne eines Elongationsmoduls**

Nach der 4'-Ppan-Transferase-katalysierten Modifikation mit dem Kofaktor 4'-Phosphopantethein interagiert die PCP-Domäne während eines katalytischen Zyklus mit mindestens drei benachbarten Domänen: der Adenylierungsdomäne sowie den *upstream* und *downstream* Kondensationsdomänen. Daneben können ggf. modifizierende Domänen vorhanden sein. Außerdem interagieren fehlbeladene PCP-Domänen mit TypII-Thioesterasen (Kapitel 1.4), die das fehlerhafte Intermediat hydrolytisch entfernen und so eine Fortsetzung der Peptidsynthese ermöglichen.

Die erste Struktur einer NRPS-PCP-Domäne, die mittels Kernspinresonanz-Spektroskopie (NMR) gelöst wurde, ist in Abbildung 1.9 gezeigt [Weber, 2000]. Wie in den bereits bekannten Strukturen von ACP-Domänen der Fettsäuresynthese wurde hier eine Struktur aus 4  $\alpha$ -Helices beobachtet („four-helix bundle“).



**Abbildung 1.9: NMR-Strukturen der PCP-Domäne des Moduls TycC3**

(A) Nach der von Weber *et al.* veröffentlichten Struktur bildet die PCP-Domäne vier  $\alpha$ -Helices aus. (B) In einer späteren NMR-Studie beschreiben Koglin *et al.* drei unterschiedliche Konformationen der PCP-Domäne.

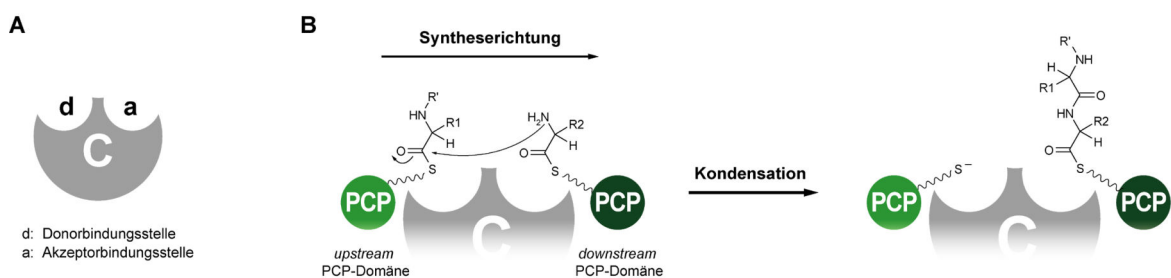
Weitergehende NMR-Untersuchungen dieser PCP-Domäne zeigten, dass die PCP-Domäne in drei voneinander verschiedenen Konformationen vorliegen kann [Koglin, 2006]. Die als A-, A/H- und H-Zustand bezeichneten Konformationen sind in Abbildung 1.9B dargestellt. Darüber hinaus konnten die Autoren für zwei natürliche Interaktionspartner nachweisen, dass die PCP-Domäne für die jeweilige Interaktion unterschiedliche Konformationen annimmt. Vermutlich ist die beobachtete Plastizität von großer Bedeutung für die vielfältigen Wechselwirkungen mit benachbarten katalytischen Domänen während der nichtribosomalen Synthese.

## Kondensationsdomäne (C-Domäne)

Kondensationsdomänen katalysieren die Ausbildung der Peptidbindungen in den Produktpeptiden, wobei die Substrate üblicherweise über Thioesterbindungen an die 4'-Ppan-Kofaktoren der PCP-Domänen gebunden sind.

Der Mechanismus, nach dem die Kondensationsdomänen die Peptidbindungsbildung katalysieren, ist noch nicht eindeutig geklärt. Es wurde aber, basierend auf den Ergebnissen biochemischer Untersuchungen, ein Modell für die Katalyse vorgeschlagen [Stachelhaus, 1998]. Diesem Modell zufolge besitzt die Kondensationsdomäne zwei Bindungsstellen, die als Donor- und Akzeptorstelle bezeichnet werden (Abbildung 1.10A). Das *upstream* PCP-gebundene Intermediat bindet in der Donorstelle der C-Domäne und das *downstream* PCP-gebundene Aminoacylsubstrat in deren Akzeptorstelle. Nach dem nukleophilen Angriff der  $\alpha$ -Aminogruppe des Aminoacylsubstrats der *downstream*-gelegenen PCP-Domäne auf den Thioester des *upstream* PCP-gebundenen Intermediats (Abbildung 1.10B, links) wird die Bindung zur *upstream* gelegenen PCP-Domäne gelöst. Das resultierende, verlängerte Peptidylintermediat ist an die *downstream* gelegene PCP-Domäne gebunden. Mit jeder folgenden Kondensationsreaktion wird das sukzessiv wachsende Peptid an das nächste Modul weitergereicht. Während also die erste Kondensationsreaktion zwischen zwei Aminoacylsubstraten erfolgt, ist das Donorsubstrat der darauffolgenden Kondensationsreaktionen ein PCP-gebundenes Peptidylintermediat. Beim Akzeptorsubstrat dagegen handelt es sich stets um ein PCP-gebundenes Aminoacylintermediat.

Anhand von Sequenzvergleichen und der biochemischen Charakterisierung von Mutanten wurde nachgewiesen, dass der zweite Histidinrest des *Core*-Motivs C3 für die Katalyse von besonderer Bedeutung ist [Bergendahl, 2002]. Man geht davon aus, dass das Histidin als katalytische Base



**Abbildung 1.10: Modellvorstellung der Kondensationsdomäne und des Mechanismus' der katalysierten Reaktion**  
(A) Die Kondensationsdomänen verfügen über zwei Bindungsstellen, über welche sie mit den Donor- und Akzeptorsubstraten interagieren. Das aktive Zentrum liegt vermutlich zwischen diesen beiden Bindungsstellen.  
(B) Nach Bindung beider Substrate erfolgt im aktiven Zentrum ein nukleophiler Angriff der deprotonierten  $\alpha$ -Amino-Gruppe des Akzeptorsubstrats auf den Thioester-Carboxylkohlenstoff des Donorsubstrats. Mit der Lösung von der *upstream* PCP-Domäne und der Bildung der Peptidbindung ist das verlängerte Peptidylintermediat auf den Kofaktor der *downstream* PCP-Domäne übertragen.

fungiert und die Aminogruppe der *downstream* PCP-gebundenen Aminosäure deprotoniert. Diese greift dann den Carboxylkohlenstoff des Thioesters der *upstream* gelegenen PCP-Domäne nukleophil an.

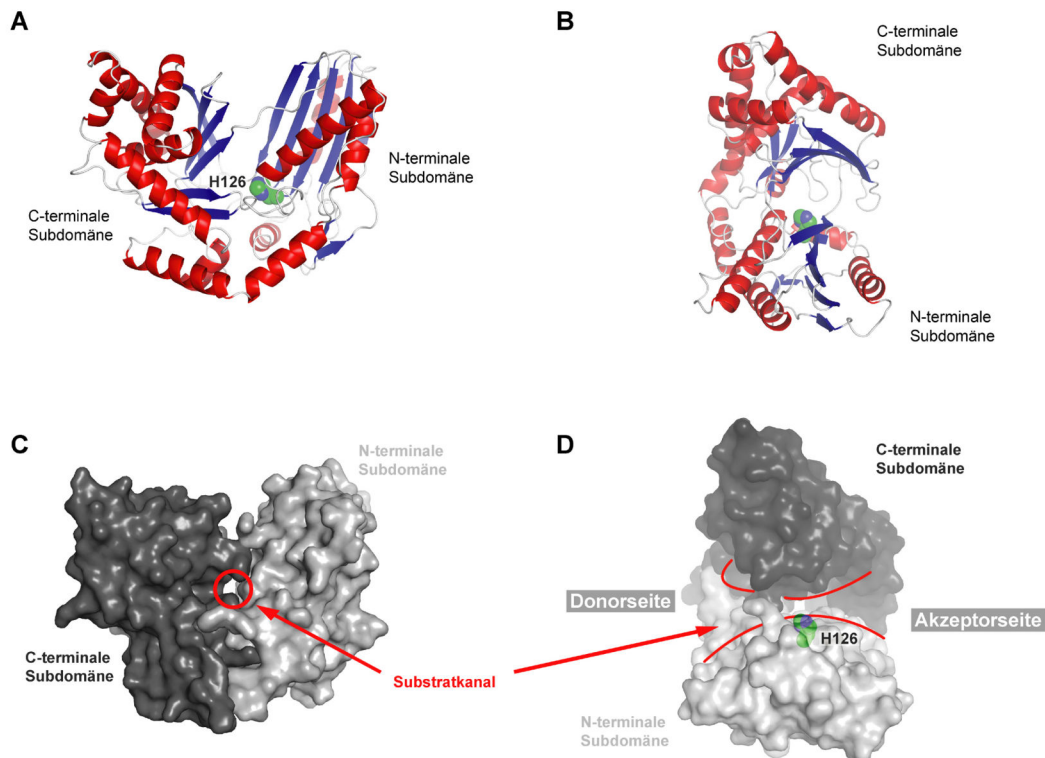
Hinsichtlich der Substratspezifität ergaben sich für die Donor- und Akzeptor-Bindungsstellen unterschiedliche Selektivitäten. Belshaw *et al.* untersuchten die Kondensationsdomäne des Moduls TycB1 auf ihre Toleranz gegenüber veränderten PCP-gebundenen Akzeptoraminosäuren [Belshaw, 1999]. In ihrer natürlichen Reaktion überträgt diese Kondensationsdomäne einen an die Synthetase TycA gebundenen D-konfigurierten Phenylalaninrest auf den L-Prolinrest der moduleigenen PCP-Domäne. Es zeigte sich, dass die Akzeptorbindungsstelle der C-Domäne eine höhere Chemo-selektivität aufweist als die Donorbindungsstelle. Auf der Donorseite wurde eine Enantio-selektivität beobachtet. Es ließ sich erstmals nachweisen, dass auch die Kondensationsdomänen eine Kontrollfunktion ausüben und so Einfluss auf die korrekte Produktzusammensetzung haben [Belshaw, 1999]. Weitere Studien zeigten, dass diejenigen C-Domänen, die auf eine Epimerisierungsdomäne folgen, für die Konfiguration des epimerisierten Intermediats spezifisch sind. Epimerisierungsdomänen katalysieren die stereochemische Umwandlung des C-terminalen, als Thioester gebundenen Aminosäurerestes, so dass sich ein Gleichgewicht zwischen Enantio- bzw. Diastereomeren einstellt. Aufgrund ihrer Selektivität der donorseitigen Bindungsstelle akzeptiert die C-Domäne nur dasjenige Intermediat, dessen C-terminaler Rest die D-Konfiguration aufweist [Clugston, 2003].

Darüber hinaus nehmen die Kondensationsdomänen Einfluss auf die Produktzusammensetzung, indem sie die 4'-Ppan-gebundenen Akzeptorsubstrate in ihrer Akzeptorbindungstasche binden. Dadurch wird nämlich verhindert, dass die Aminoacyl-Zwischenstufen interner Module die Synthese initiieren, was zwangsläufig zu verkürzten Produkten falscher Zusammensetzung führen würde. Der Nachweis dieser Kontrollfunktion wurde in Studien mit Elongationsmodulen erbracht. Diese sind üblicherweise nicht in der Lage, die nichtribosomale Peptidsynthese zu initiieren. Nach Entfernung der *upstream* gelegenen Domänen war das verbleibende A-PCP-Modul dazu jedoch befähigt. Anhand der Untersuchung verschiedener rekombinanter Proteine, die sich in der Anzahl der entfernten Domänen unterschieden, konnte dieser Effekt eindeutig auf die C-Domäne zurückgeführt werden [Linne, 2000].

Im Jahr 2002 veröffentlichten Keating *et al.* die erste Struktur einer NRPS-Kondensationsdomäne, des VibH. Wie in Abbildung 1.11A dargestellt, besteht diese aus zwei V-förmig angeordneten Untereinheiten, die ihrerseits Ähnlichkeit zur Chloramphenicol-Acetyltransferase [Lewendon, 1994] und Dihydrolipoamid-Acetyltransferase [De Crecy-Lagard, 1995] aufweisen. Zwischen diesen N- und C-terminalen Subdomänen, die in Abbildung 1.11B unterschiedlich schattiert sind,

## 1 Einleitung

verläuft ein Kanal, der das Enzym durchquert und in dessen Mitte das aktive Zentrum der C-Domäne liegt. Der Substratkanal mit seinen donor- und akzeptorseitigen Zugängen sowie die Lage des katalytisch aktiven Histidinrestes H126 sind in Abbildung 1.11C durch rote Linien bzw. als Kalottenmodell der Histidin-Seitenkette angedeutet.



**Abbildung 1.11: Kristallstruktur der isoliert stehenden Kondensationsdomäne VibH [Keating, 2002]**

(A) Die Struktur von VibH mit dem konservierten Histidinrest H126 im aktiven Zentrum der Kondensationsdomäne. (B) Aufsicht auf die Kondensationsdomäne VibH mit der Seitenkette des Restes H126 als Kalottenmodell dargestellt. (C) Perspektive wie in (A). In der Oberflächendarstellung von VibH wird der Kanal sichtbar, der von beiden Seiten des Enzyms zugänglich ist und in dessen Mitte das aktive Zentrum liegt. (D) Perspektive wie in (B). Der Verlauf des Substratkanals ist durch rote Linien hervorgehoben, die Position des H126 im Kanal ist angedeutet.

VibH stellt in zweierlei Hinsicht einen Sonderfall dar. Zum einen empfängt VibH das Donorsubstrat *in trans*, d. h. an den Kofaktor einer anderen Synthetase, VibB, gebunden. Das Akzeptorsubstrat ist gar nicht enzymgebunden und bindet ohne Unterstützung auf der Akzeptorseite. Daher dürften sich die Wechselwirkungen zur Substraterkennung von denen des allgemeineren Falls, bei dem zwei Kofaktor-gebundene Substrate *in cis* vorliegen, unterscheiden. Zum anderen stellte sich bei der biochemischen Charakterisierung heraus, dass das Protein nach Mutation des vermeintlich essentiellen Histidinrestes (H126A) noch eine messbare Restaktivität von etwa 10 % aufweist. Dies steht im Gegensatz zu den bisherigen Beobachtungen mit den

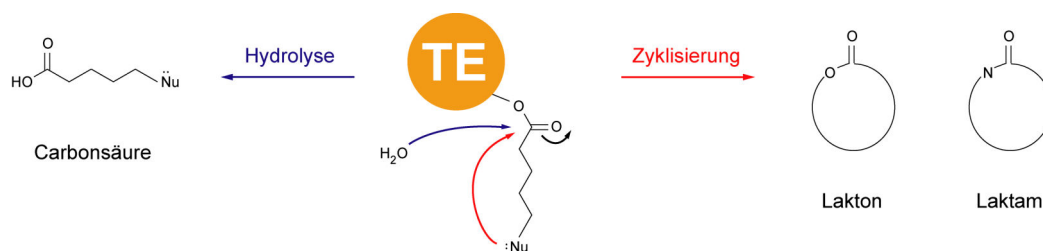
Kondensationsdomänen der Module TycB1, EntF und TycC6, nach denen keine Aktivität zu erwarten gewesen wäre [Bergendahl, 2002; Roche, 2003; Schönafinger, 2003].

Interessant ist auch, dass kürzlich nachgewiesen wurde, dass die Kondensationsdomäne SgcC5 aus dem Synthetasecluster des Endiin-Antibiotikums C-1027 neben ihrer natürlichen Reaktion eine weitere ligierende Funktion zeigt. Während die Domäne normalerweise eine Peptidbindung knüpft, konnte diese C-Domäne mit entsprechend veränderten Substraten auch Esterbindungen bilden [Lin, 2009]. Dies verdeutlicht noch einmal die Universalität und das Potential dieser Enzyme. Um die Möglichkeiten, diese Katalysatoren an neue Reaktionen anzupassen, voll auszuschöpfen, müssen jedoch erst die Faktoren identifiziert werden, die die Substraterkennung und damit die Substratspezifität bestimmen.

## 1.4 Optionale und modifizierende Domänen

### Thioesterasedomäne (TE-Domäne)

Die Abspaltung von der Synthetase stellt häufig den letzten Schritt der Synthese dar. Dieser wird in den meisten Fällen von einer Thioesterasedomäne katalysiert, die sich am C-Terminus des Synthetaseclusters befindet. In einigen Clustern bewirken TE-Domänen die Hydrolyse des linearen, PCP-gebundenen Peptids, wie dies beim Penicillin-Vorläufer ACV oder beim Vancomycin beobachtet wird. Weiter verbreitet dagegen ist die TE-katalysierte Makrozyklisierung, welche zyklische oder verzweigt-zyklische Makrolaktone und Makrolaktame hervorbringt. So wird beispielsweise die Makrozyklisierung der Makrolaktone Surfactin, Fengycin und Daptomycin bzw. die der Makrolaktame Gramicidin und Tyrocidin von den Thioesterasedomänen ihrer jeweiligen Synthetasecluster katalysiert.



**Abbildung 1.12: Reaktionsprodukte Thioesterase-katalysierter Domänen**

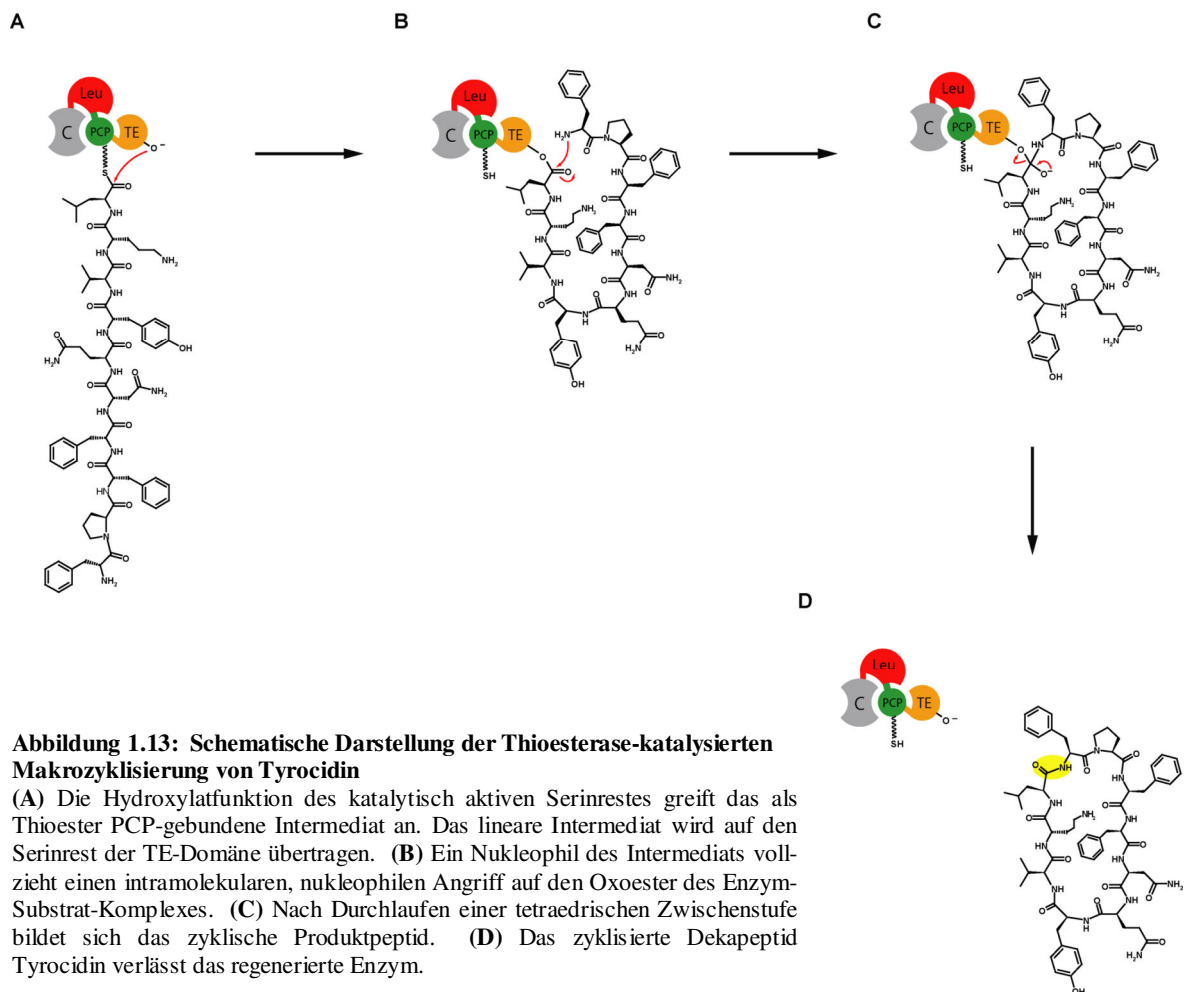
Neben der Zyklisierung, die meist Makrolaktone und Makrolaktame liefert, besteht auch die Möglichkeit der Hydrolyse. In letzteren Fall ist das Produkt die freie Carbonsäure der Peptidverbindung. Während dies in den Biosyntheseclustern des Vancomycin sowie der Penicillin- und Cephalosporin-Vorläufer der reguläre Weg der Produktfreisetzung ist, kann sie bei zyklisierenden Thioesterasedomänen auch als Nebenreaktion auftreten.

## 1 Einleitung

Allerdings ist die Thioesterase nicht die einzige Domäne, die die Abspaltung des Produkts von der Synthetase katalysieren kann. Im Falle der Cyclosporin A-Synthetase aus *Tolyplocadium niveum* katalysiert eine C-terminale Kondensationsdomäne die Makrolaktamisierung des Produkts Cyclosporin A [Weber, 1994]. Daneben besteht aber auch die Möglichkeit, dass eine Reduktasedomäne das Syntheseprodukt als Aldehyd abspaltet, wie es beim linearen Gramicidin beobachtet wird [Kessler, 2004].

Die Thioesterase-katalysierte Zyklisierung verläuft nach einem den Serinproteasen ähnlichen Mechanismus. Daher verfügen auch die Thioesterasedomänen für ihre biologische Funktion über eine katalytische Triade, die meist einen Aspartat-, einen Histidin- und einen Serinrest umfasst.

Der Mechanismus der Reaktion ist in Abbildung 1.13 am Beispiel des Tyrocidins gezeigt. Es erfolgt zunächst die Übertragung des linearen Dekapeptids vom Kofaktor der PCP-Domäne auf den katalytisch aktiven Serinrest. Der Oxoester des kovalenten Enzym-Acyl-Intermediats wird durch ein intramolekulares Nukleophil angegriffen, was zunächst zu einer tetraedrischen Zwischenstufe führt. Unter Bildung des zyklischen Produkts spaltet sich das Peptid von der Synthetase ab.





In den Biosynthesen des Gramacidins sowie der Siderophore Bacillibactin und Enterobactin zeigt die Thioesterasedomäne eine zusätzliche Aktivität. In diesen iterativen Thioesterasedomänen werden mehrere vom Cluster synthetisierte Peptidylintermediate *head-to-tail* verknüpft und anschließend zyklisiert. Auf diese Weise entstehen beispielsweise Dimere (Gramicidin) oder Trimere (Bacillibactin, Enterobactin) der gebildeten Peptidylintermediate [Mootz, 2002].

Wie die Adenylierungsdomänen, sind die meisten Thioesterasedomänen auch unabhängig von ihren *upstream* gelegenen Synthetasen katalytisch aktiv. Die isolierte Thioesterasedomäne des Tyrocidin Synthetaseclusters katalysiert die Zyklisierung eines Tyrocidin-analogen, löslichen Substrats, das als Thioester des S-N-Acetylcysteamin (SNAC) aktiviert ist [Trauger, 2000]. Dieses SNAC „imitiert“ die Thioester-nahe Umgebung des PCP-gebundenen 4'-Phosphopantethein-Kofaktors im natürlichen Substrat. Darüber hinaus ließ sich mit der Tyrocidin TE-Domäne auch Gramicidin durch Dimerisierung und Zyklisierung von Pentapeptideinheiten herstellen. Die „Abkopplung“ der TE-Domäne ermöglichte ihre genaue Charakterisierung. Unter Einsatz synthetisch veränderter Substratpeptide konnte die Toleranz gegenüber nicht-natürlichen Substraten untersucht werden.

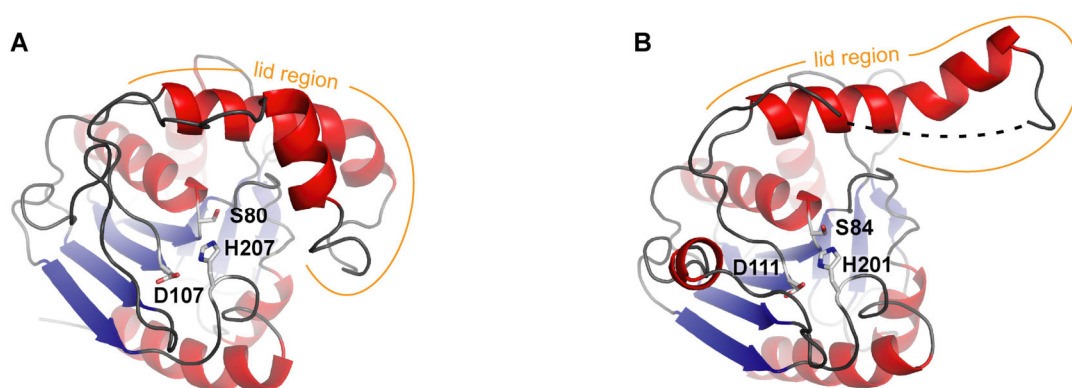
Nachfolgende Studien mit Thioesterasedomänen verschiedener Synthesecuster ergaben, dass die TE-Domänen nur eine begrenzte Zahl von Aminosäureresten in ihren Substraten erkennen und Veränderungen in den meisten Positionen tolerieren. In den von den TE-Domänen erkannten Positionen sind allerdings sowohl die Art des Aminosäurerestes als auch dessen stereochemische Konfiguration für die Zyklisierung wichtig. Die Bedeutung der Reste an diesen Positionen konnte auch für andere TE-Domänen gezeigt werden, da mit entsprechend abweichenden Substratpeptiden keine Zyklisierung beobachtet wurde [Kohli, 2001; Sieber, 2003]. Einige Thioesterasedomänen, die solche SNAC-aktivierten Substrate nicht akzeptieren, ließen sich mit Hilfe der stärker aktivierten Thiophenyl-Substrate untersuchen [Sieber, 2004].

Es fiel außerdem auf, dass mit den isolierten TE-Domänen höhere Anteile von hydrolysiertem, linearem Produkt erhalten werden. Dies ist vermutlich darauf zurückzuführen, dass durch die Entfernung der benachbarten Domänen der Zugang für Wasser zum aktiven Zentrum erleichtert wird. Um die Hydrolyse des TE-gebundenen Substrats zu vermeiden, ist es wichtig, Wasser aus den aktiven Zentren der Thioesterase fernzuhalten. Es konnte gezeigt werden, dass sich die Hydrolyse als Nebenreaktion durch die Verwendung ionischer Flüssigkeiten oder in Gegenwart nichtionischer Detergentien zurückdrängen lässt [Yeh, 2004; Wagner, 2006].

## 1 Einleitung

Auf die Bedeutung der Zyklisierung für die biologische Aktivität der Verbindungen wurde bereits in Kapitel 1.1 hingewiesen. Durch die Makrozyklisierung wird das Produktpeptid in seiner Beweglichkeit eingeschränkt und in seiner biologisch aktiven Konformation gehalten.

Wie die strukturelle Charakterisierung der NRPS-Thioesterasedomänen ergab, lassen sie sich dem Faltungstyp der  $\alpha/\beta$ -Hydrolasen zuordnen [Bruner, 2002; Samel, 2004]. In den TE-Domänen wurde eine Insertion beobachtet, die sich an den  $\beta$ -Strang  $\beta_6$  anschließt und das aktive Zentrum wie ein Deckel verschließen kann. Daher wurde diese Insertion in Analogie zu den Lipasen als *lid region* bezeichnet. Abbildung 1.14 zeigt die Kristallstrukturen der Surfactin- und Fengycin TE-Domänen. Die *lid regions* und die Reste der katalytischen Triaden sind hervorgehoben.



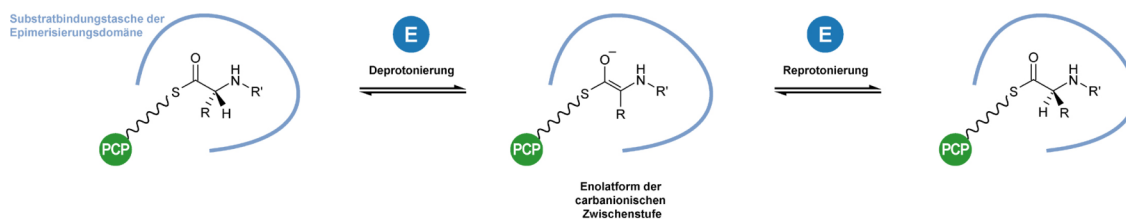
**Abbildung 1.14: Kristallstrukturen der Surfactin- und Fengycin Thioesterasedomänen**

Die TE-Domänen enthalten neben dem  $\alpha/\beta$ -Hydrolasen-Faltungstyp eine *lid region*, wie sie in Lipasen beobachtet wird. (A) Die Surfactin TE (SrfTE) ist die erste strukturell charakterisierte NRPS-Thioesterasedomäne. In der Kristallstruktur wurden zwei Proteinmoleküle pro asymmetrischer Einheitszelle beobachtet, die sich hauptsächlich in der Konformation ihrer *lid regions* unterscheiden [Bruner, 2002]. (B) In der Struktur der Fengycin TE-Domäne (FenTE) ist ein Teil der *lid region* in der Elektronendichte nicht definiert. Die Enden dieses Bereichs sind durch eine unterbrochene Linie miteinander verbunden.

Über die strukturellen Grundlagen der Substratspezifität ist bislang wenig bekannt. Basierend auf Arbeiten mit der Surfactin TE-Domäne wurden zwei Bindungsmodelle vorgeschlagen [Bruner, 2002; Tseng, 2002]. Das erste Modell beschreibt eine Bindung des linearen Lipopeptids in einem wannenförmigen aktiven Zentrum. Das zweite Modell geht davon aus, dass nur die an der Zyklisierung beteiligten Reste des Peptids mit der TE-Domäne interagieren, während der größte Teil aus dem aktiven Zentrum herausragt.

## Epimerisierungsdomäne (E-Domäne)

Die Epimerisierungsdomänen liegen direkt hinter den PCP-Domänen ihres Moduls und katalysieren die chemische Umwandlung zwischen der L-konfigurierten und der D-konfigurierten Form der thioestergebundenen Aminosäure des Aminoacyl- bzw. Peptidylintermediats. Dadurch wird die strukturelle Vielfalt der Produktpeptide enorm bereichert.



In den Epimerisierungsdomänen enthaltenden Modulen aktiviert die zugehörige A-Domäne spezifisch die L-Aminosäure und überträgt sie auf den 4'-Ppan-Kofaktor der PCP-Domäne. Während in Initiationsmodulen die PCP-gebundene Aminosäure das Substrat der Epimerisierungsdomäne darstellt, findet in Elongationsmodulen zuerst die Übertragung des *upstream* synthetisierten Intermediats statt. Erst dann, auf der Stufe des PCP-gebundenen Peptidylintermediats, erfolgt die Epimerisierung am C-terminalen Aminosäurerest.

Weitere Untersuchungen an Epimerisierungsdomänen haben ergeben, dass die Reaktion in beide Richtungen katalysiert wird und so ein Gleichgewicht zwischen den Enantiomeren- bzw. Diastereomeren entsteht [Stachelhaus, 2000]. Aufgrund der Selektivität der nachfolgenden C-Domäne für die D-Konfiguration der C-terminalen Aminosäure des enzymgebundenen Donorsubstrats findet sich an der entsprechenden Stelle im Produktpeptid nur die D-konfigurierte Aminosäure wieder [Clugston, 2003].

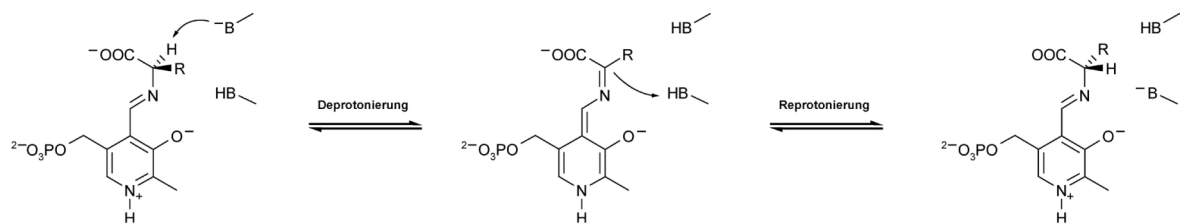
Ein Beispiel für eine Alternative zur *in situ*-Epimerisierung von L-konfigurierten Aminosäuren findet sich in der Cyclosporin A-Synthetase aus *Tolyplocadium niveum*. Die Adenylierungsdomäne aktiviert dort spezifisch D-Alanin, welches durch eine separate Epimerase, die ebenfalls durch das Cyclosporin Synthetase Operon kodiert ist, bereitgestellt wird [Hoffmann, 1994]. Ebenso wurde über eine Variante von Kondensationsdomänen berichtet, die zusätzlich eine Epimerisierung katalysiert. Diese dualen Kondensations- und Epimerisierungsdomänen werden später in diesem Kapitel vorgestellt. Schließlich wurde auch in der Pyochelin Synthetase die Epimerisierung eines Cysteinrestes vor dessen Heterozyklisierung zu einem Thiazolin beobachtet. Die Epimerisierungsaktivität wurde 310 bis 360 Aminosäuren großen Insertionen in den Adenylierungsdomänen der entsprechenden Module zugeschrieben [Patel, 2003].

## 1 Einleitung

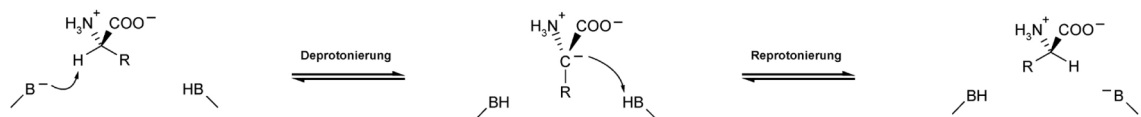
Bei den Aminosäureepimerasen kann generell zwischen Kofaktor-abhängigen und Kofaktor-unabhängigen Epimerasen unterschieden werden. Eine Klasse von Epimerasen verwendet den Kofaktor 5-Pyridoxalphosphat (PLP), die andere Klasse katalysiert die Epimerisierung unabhängig von Kofaktoren. Bei der PLP-abhängigen Epimerisierung werden die Aminosäuren aktiviert, indem ihre Aminogruppe kovalent an den Kofaktor gebunden wird. Damit wird unter Bildung eines Aldimins die Deprotonierung und damit die Ausbildung einer carbanionischen Zwischenstufe erleichtert.

Die Kofaktor-unabhängigen Epimerasen katalysieren die Epimerisierung der freien, nicht-aktivierten Aminosäuren. Diese Enzymklasse enthält im aktiven Zentrum meist zwei Cysteinreste, die als katalytische Base und katalytische Säure wirken und die De- bzw. Reprotonierung des Substrats herbeiführen [Tanner, 2002]. Die Mechanismen der beiden Enzymklassen zeigt Abbildung 1.15.

### 5'-Pyridoxalphosphat-abhängige Epimerisierung



### Kofaktor-unabhängige Epimerisierung



### Abbildung 1.15: Mechanismen der Kofaktor-abhängigen bzw. -unabhängigen Aminosäure-Epimerisierung

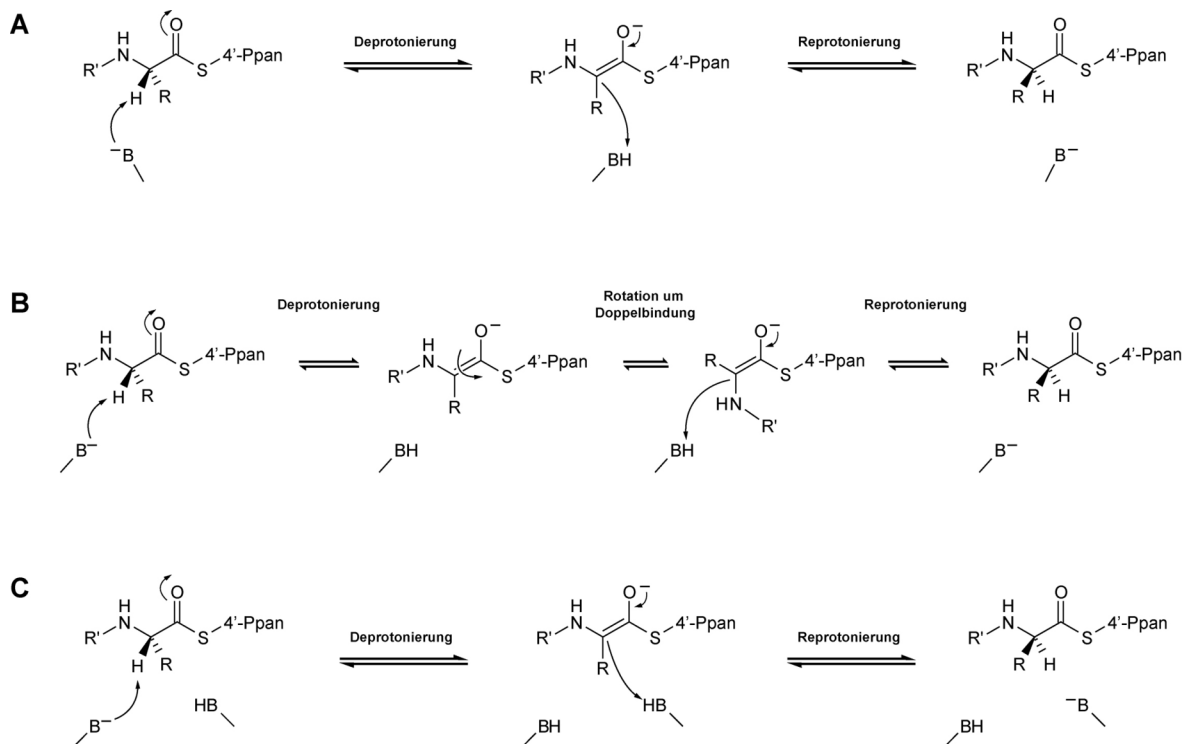
In der Kofaktor-abhängigen Epimerisierung wird die Aminosäure kovalent an 5'-Pyridoxalphosphat gebunden. Die Ausbildung eines Aldimins erhöht die Azidität in der  $C_\alpha$ -Position. Die Kofaktor-unabhängige Epimerisierung wird von zwei Cysteinresten katalysiert, die als katalytisches Säure-Base-Paar agieren.

Die Epimerisierungsdomänen aus nichtribosomalen Peptidsynthetasen lassen sich keiner der beiden Gruppen zuordnen. Den bisherigen Erkenntnissen zufolge benötigen sie keinen Kofaktor. Sie zeigen aber auch keinerlei Ähnlichkeit zur oben genannten Kofaktor-unabhängigen Gruppe. Stattdessen legen phylogenetische Untersuchungen eine Verwandtschaft zu den NRPS-Kondensationsdomänen nahe, und Strukturvorhersagen zufolge besteht zu diesen auch eine strukturelle Ähnlichkeit. In biochemischen Studien wurde die Substratspezifität von NRPS-Epimerisierungsdomänen untersucht [Linne, 2001; Stein, 2005] und die Frage geklärt, ob es

spezifische Unterschiede zwischen Aminoacyl- und Peptidyl-epimerisierenden Epimerisierungsdomänen gibt, oder ob sie sich nur durch die jeweilige Position im Syntheseprozess unterscheiden. Es konnte nachgewiesen werden, dass die Epimerisierungsdomänen auf ihre jeweiligen Substrate spezialisiert sind [Stein, 2006].

Mit Hilfe von Mutationsstudien wurde versucht, die am Mechanismus beteiligten Aminosäurereste ausfindig zu machen. Allerdings ist der katalytische Mechanismus der Epimerisierungsdomänen noch nicht eindeutig geklärt. Es wird davon ausgegangen, dass im Verlauf der Reaktion das  $C_{\alpha}$ -Proton von einer katalytischen Base abstrahiert wird [Stachelhaus, 2000]. Die anionische Zwischenstufe wird als Enolat stabilisiert, da in den 4'-Ppan-gebundenen Substraten praktisch keine mesomere Stabilisierung durch die Thioesterfunktion erfolgt. Die Reprotonierung aus dem anderen enantiotopen bzw. diastereotopen Halbraum führt zum epimerisierten Produkt.

Es ist auch noch ungeklärt, ob die Reaktion über einen Ein-Basen- oder einen Zwei-Basen-Mechanismus katalysiert wird. Diese beiden Mechanismen sind in Abbildung 1.16 gegenübergestellt. Im ersten Fall werden beide Teilschritte, die Deprotonierung wie auch die Reprotonierung, von einer einzigen Base katalysiert. Im Laufe dieser Reaktion muss sich das Intermediat um die



**Abbildung 1.16: Vergleichende schematische Darstellung des Ein-Basen- und Zwei-Basen-Mechanismus**

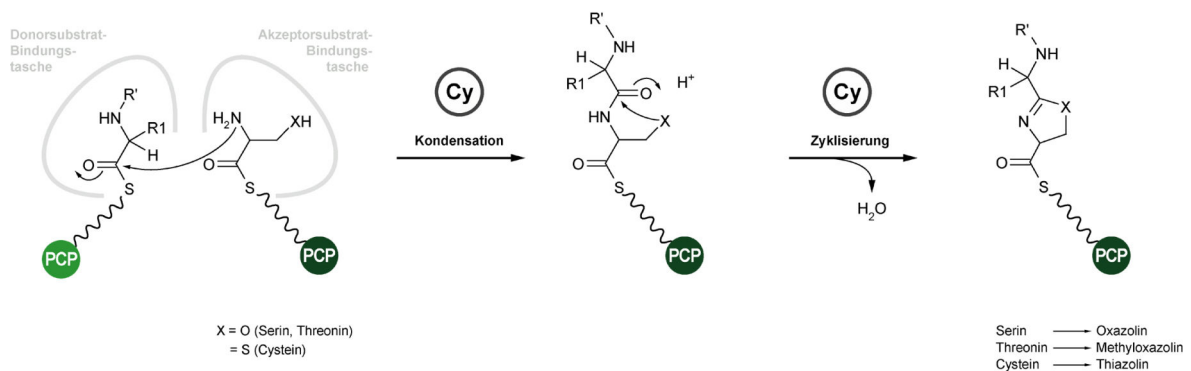
(A) Nach Deprotonierung des Substrats wird das Enolat von derselben Base aus dem anderen enantio-/diastereotopen Halbraum reprotoniert. (B) Alternativ zu (A) kann eine statische Base vorliegen. In diesem Fall muss im Enolat eine Rotation um die Doppelbindung erfolgen. (C) Ein Proton des Substrats wird durch eine katalytische Base abstrahiert. Die Enolatzwischenstufe wird anschließend von einer anderen, protoniert vorliegenden Base reprotoniert.

## 1 Einleitung

Doppelbindung des Enolats drehen, so dass nach der Reprotonierung eine veränderte Konfiguration erhalten wird. Im zweiten Fall abstrahiert eine Base ein Proton des Substrats, während eine andere protonierte Base die anionische Zwischenstufe aus dem enantio- bzw. diastereotopen Halbraum reprotoniert. Dieser Fall entspricht dem bei den Kofaktor-unabhängigen Epimerasen beobachteten Mechanismus.

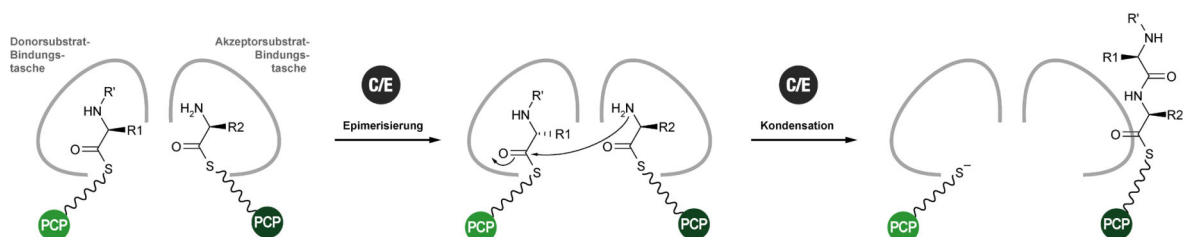
### Zyklisierungsdomäne (Cy-Domäne)

Die Zyklisierungsdomäne ist eng mit der Kondensationsdomäne verwandt und kann diese ersetzen. Neben der Verlängerung um einen Serin-, Threonin- oder Cysteinrest katalysiert die Zyklisierungsdomäne zusätzlich den nukleophilen Angriff der Seitenkette auf die Carboxylgruppe der vorausgehenden Peptidbindung sowie die anschließende Dehydratisierung des neugebildeten, fünfgliedrigen Heterozyklus'. So entstehen Oxazoline, Methyloxazoline bzw. Thiazoline, die gegebenenfalls durch Oxidations- oder Reduktionsdomänen weiter modifiziert werden können.



### Duale Kondensations- und Epimerisierungsdomänen (C/E-Domäne)

Duale Kondensations- und Epimerisierungsdomänen katalysieren sowohl eine Kondensationsreaktion als auch eine Epimerisierungsreaktion. Die Existenz dieser bifunktionellen Domänen war im Arthrophactin Biosynthesecoluster vermutet worden, da im Produktpeptid D-konfigurierte Aminosäuren auftreten, aber weder Epimerisierungsdomänen, noch D-Aminosäure-spezifischen



Adenylierungsdomänen gefunden wurden. Die dualen C/E-Domänen können anhand ihrer Proteinsequenz von den konventionellen Kondensationsdomänen unterschieden werden [Balibar, 2005]. Über diese Sequenzunterschiede ließen sich C/E-Domänen auch in den Synthetasen der Syringomycin-, Syringopeptin- und Ramoplanin-Synthesecluster nachweisen.

Interessanterweise wird in diesen Domänen die C-terminale Aminosäure des *upstream* gebundenen Intermediats epimerisiert. Diese Reaktion erfolgt jedoch nur dann, wenn auf der Akzeptorseite ein PCP-gebundenes Aminoacylintermediat vorliegt.

### Weitere katalytische Domänen in der nichtribosomalen Peptidsynthese

Oxidationsdomänen (Ox-Domänen), die häufig in Verbindung mit Zyklisierungsdomänen angetroffen werden, katalysieren die FMN-abhängige Reaktion der Oxa-/Thiazolin-Zwischenprodukte zu den aromatischen Oxa-/Thiazolen. Sie wurden an zwei verschiedenen Positionen beobachtet, zum einen direkt hinter PCP-Domänen gelegen, zum anderen insertiert in Adenylierungsdomänen.

Dagegen wirken die Reduktasedomänen (Red-Domänen) auf Oxa-/Thiazoline und liefern die vollständig reduzierten Oxa-/Thiazolidine. Die Reduktasedomänen benötigen für diese Reaktion den Kofaktor NADPH.

Die *N*-Methylierung ist eine weitere Art der Modifizierung, die das Peptidprodukt vor Proteolyse schützt. Die Methylgruppen werden in einer von *N*-Methylierungsdomänen (Mt-Domänen) katalysierten Reaktion von dem Kosubstrat *S*-Adenosylmethionin (SAM) auf das Intermediat übertragen.

Schließlich gibt es noch weitere Enzyme, die spezielle, nichtproteinogene Aminosäuren synthetisieren, die *in trans* auf PCP-gebundene Zwischenstufen einwirken oder die die abgespaltenen Produkte nichtribosomaler Peptidsynthetasen modifizieren.

So wird beispielsweise das Vancomycin als lineares Heptapeptid von seiner Synthetase abgespalten. Erst dann schließen sich die oxidativen Kupplungsreaktionen an, in denen die für die Konformation und für die biologische Aktivität essentiellen drei Makrozyklen entstehen. Des Weiteren werden durch spezielle Glykosyltransferasen die Zuckereinheiten auf das Peptid übertragen, die ebenfalls für die Aktivität dieses Antibiotikums unerlässlich sind. Der Syntheseprozess des Antibiotikums Vancomycin ist ein anschauliches Beispiel für die Modifikationsvielfalt. Ein umfassendes Bild über den Syntheseprozess des Antibiotikums Vancomycin als Beispiel für die Modifikationsvielfalt geben Hubbard und Walsh [Hubbard, 2003].

## 1.5 Fettsäuresynthasen und Polyketidsynthasen

Die Biosynthese der Fettsäuren (FAS) und der Polyketide (PKS) verläuft nach einem ähnlichen Mechanismus wie die nichtribosomale Peptidsynthese [Fischbach, 2006; Smith, 2007]. In beiden Fällen unterscheidet man prinzipiell zwischen zwei Organisationen der Synthasen, den multifunktionalen Enzymen (Typ I) einerseits und den löslichen, monofunktionalen Enzymen (Typ II) andererseits. Außerdem gibt es Ähnlichkeiten sowohl in der Verwendung eines Thiotemplat-Mechanismus' als auch in der Art der katalysierten, chemischen Reaktionen und im modularen Aufbau der Synthetasen. Beide Synthesen können repetitiv erfolgen, indem die Intermediate die Enzyme in mehreren Zyklen durchlaufen. Für die Polyketidsynthese existiert jedoch auch, wie in vielen nichtribosomalen Peptidsynthetasen, ein nicht-repetitiver, linearer Synthesemodus.

Die Polyketidsynthasen (PKS) verwenden als Substrate Malonyl- und Methylmalonyl-CoA, die von den Acetyltransferase (AT)-Domänen ausgewählt werden. In der nicht-repetitiven Synthese überträgt jede AT-Domäne den jeweiligen CoA-gebundenen Acylrest auf den 4'-Ppan-Kofaktor der zugehörigen *Acyl Carrier Protein* (ACP)-Domäne. Die Knüpfung der C-C-Bindungen wird von den Ketosynthase (KS)-Domänen katalysiert. Dabei werden die sukzessiv verlängerten Intermediate von den ACP-Domänen auf einen konservierten, katalytisch aktiven Cysteinrest der KS-Domäne übertragen. Nach dem Mechanismus der *Claisen*-Kondensation erfolgen die Decarboxylierung des *downstream* ACP-gebundenen Substrats und der nukleophile Angriff auf den Thioester des KS-gebundenen Intermediats. Das verlängerte Intermediat ist damit an die *downstream* gelegene ACP-Domäne gebunden. In Analogie zur oben dargestellten nichtribosomalen Peptidsynthese wird das Intermediat sukzessive weitergereicht und schließlich von einer Thioesterase unter Makrozyklisierung abgespalten. Die in der Fettsäuresynthese (FAS) vorkommenden Ketoreduktase (KR)-, Dehydratase (DH)- und Enoylreduktase (ER)-Domänen existieren ebenfalls in Polyketidsynthasen. Allerdings können eine oder mehrere dieser Domänen in einigen Modulen der Polyketidsynthasen fehlen, was sich im Produkt widerspiegelt, da an der entsprechenden Position die Kohlenstoffkette nicht vollständig reduziert wird und eine funktionelle Gruppe erhalten bleibt. Ähnlich wie bei der nichtribosomalen Peptidsynthese entstehen auf diese Weise strukturell wie funktionell vielseitige Naturstoffe mit pharmakologisch wertvollen Eigenschaften.

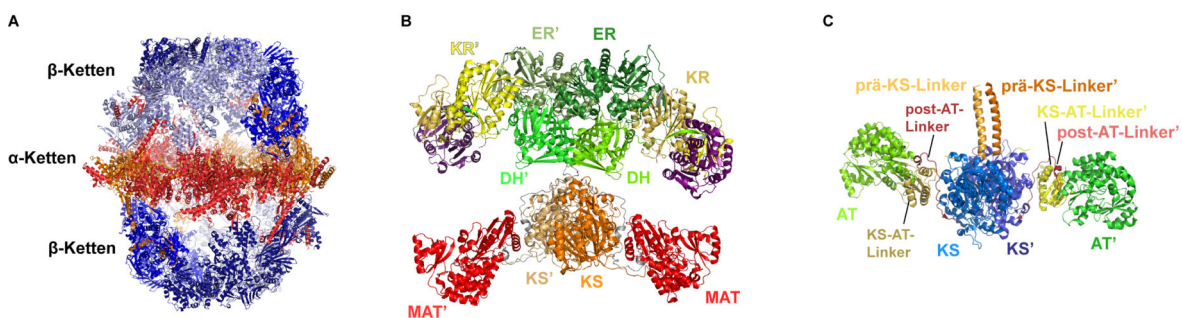
Die vergangenen Jahre brachten neue Einblicke in die strukturelle Organisation von Fettsäuresynthasen und Polyketidsynthasen. Trotz der Gemeinsamkeiten mit den nichtribosomalen Peptidsynthetasen (modularer Aufbau, Thiotemplat-Mechanismus) lassen sich die strukturellen



Erkenntnisse nur begrenzt auf die Proteine der nichtribosomalen Peptidsynthetasen übertragen, da sowohl die Fettsäuresynthetasen als auch die Polyketidsynthetasen höhere Oligomerisierungszustände aufweisen.

Auf dem Gebiet der Typ I-Fettsäuresynthase-Komplexe lieferte die Arbeitsgruppe Prof. Ban einen beträchtlichen Anteil der gewonnenen Erkenntnisse. Zwar erlaubten die Kristallstrukturen der dimeren FAS aus dem Schwein zunächst nur eine Platzierung der Domänen, spätere Strukturen gaben aber auch Aufschluss über die Domänen-verbindenden Linker. Die Struktur des 2.6 MDa großen, heterododekameren FAS-Komplexes aus *Thermomyces lanuginosus* [Jenni, 2007] gab ebenfalls Einblicke auf molekularer Ebene.

Schließlich zeigten sich in den Kristallstrukturen der KS-AT-Fragmente der Module 3 und 5 der 6-Deoxyerythronolid Synthase (DEBS) die Wechselwirkungen zwischen den Domänen einerseits und zwischen den Monomeren dieser als Dimere vorliegenden Synthetasen andererseits. Derartige Interaktionen werden bei den monomeren nichtribosomalen Peptidsynthetasen nicht beobachtet.



**Abbildung 1.17: Kristallstrukturen zweier Fettsäuresynthetasen und einer Bidomäne einer Polyketidsynthase**

(A) In der Fettsäuresynthase aus *Thermomyces lanuginosus* sind die katalytischen Einheiten auf zwei Polypeptidketten verteilt. Der 2.6 MDa große Komplex setzt sich aus 6  $\alpha$ - und 6  $\beta$ -Ketten zusammen [Maier, 2008]. (B) In der Fettsäuresynthase des Schweins sind alle katalytischen Domänen auf einer Peptidkette enthalten. Die aktive Form dieses Enzyms ist das dargestellte Dimer, das sich in einen ‚kondensierenden Flügel‘ (unten, KS- und AT-Domäne) sowie einen ‚modifizierenden Flügel‘ (oben, DH- ER-, KR-Domänen) unterteilen lässt. (C) In der Struktur des Ketosynthase-Acyltransferase-Bidomänenfragments aus dem 5. Modul der 6-Deoxyerythronolid-Synthase (DEBS) zeigt sich analog zur Säugetier-Fettsäuresynthase eine Dimerisierung über die Ketosynthase-Domänen [Tang, 2006]. Außerdem sind drei Linkerbereiche (prä-KS-; KS-AT-; post-AT-Linker) definiert.

Die dimere Quartärstruktur der Polyketidsynthetasen zeigt sich auch in sämtlichen, in der Proteindatenbank (<http://www.rcsb.org/>, Stand: Mai 2009) deponierten Kristallstrukturen PKS-abgeleiteter Synthetasefragmente. So lagen die Thioesterasedomänen der 6-Deoxyerythronolid- und Pikromycin-Biosynthesecluster im Kristallgitter dimer vor, wobei die Dimerisierung über zwei N-terminale  $\alpha$ -Helices eines jeden Monomers erfolgte. Ansonsten zeigten die beiden Thioesterasedomänen große strukturelle Ähnlichkeit zu den beiden strukturell charakterisierten NRPS-Thioesterasedomänen Surfactin TE und Fengycin TE [Samel, 2004].

## **Aufgabenstellung der vorliegenden Arbeit**

Mit der vorliegenden Arbeit sollten unter Verwendung röntgenkristallographischer Methoden die Strukturen einzelner Domänen sowie mehrere Domänen umfassender Proteine nichtribosomaler Peptidsynthetasen untersucht werden.

Ein Schwerpunkt lag auf der Untersuchung der Kondensationsdomänen, die trotz ihrer zentralen Rolle in der nichtribosomalen Peptidsynthese biochemisch und strukturell noch wenig untersucht worden waren. Zu Beginn der Arbeit war lediglich die Röntgenstruktur einer Kondensationsdomäne, der VibH, veröffentlicht. Da diese aber, wie in Kapitel 1.3 erläutert, eine Sonderstellung einnimmt, sollte geklärt werden, ob die Struktur repräsentativ für Kondensationsdomänen ist. Daher sollten strukturelle Untersuchungen an anderen Kondensationsdomänen dazu beitragen, ein genaueres Bild gemeinsamer struktureller Gegebenheiten in Kondensationsdomänen zu erhalten.

Ein weiterer Schwerpunkt wurde auf die strukturelle Aufklärung von Epimerisierungsdomänen gelegt. Das Interesse für diesen Domärentyp liegt zum einen darin, dass sich die NRPS-Epimerisierungsdomänen keiner der bekannten Gruppen Kofaktor-unabhängiger Aminosäure-Racemasen zuordnen lassen. Voraussichtlich liegt hier ein neuer Katalysemechanismus vor. Zum anderen stellen Epimerisierungsdomänen häufig die C-terminale Domäne ihrer Synthetase dar (vgl. Abbildung 1.3). Damit kommt ihnen auch eine wichtige Rolle bei der gegenseitigen Wechselwirkung und Erkennung mit der nachfolgenden Peptidsynthetase zu. Epimerisierungsdomänen, für die eine Verwandtschaft mit den Kondensationsdomänen vorhergesagt wurde, waren bislang strukturell nicht charakterisiert.

Außerdem sollten die früheren Arbeiten mit der Fengycin Thioesterase fortgeführt werden. Nach der Lösung der Kristallstruktur in der Diplomarbeit sollten Informationen über einen Enzym-Substrat-Komplex und damit Einblick in die entscheidenden und spezifitätsbestimmenden Wechselwirkungen erhalten werden.

## 2 Material

### 2.1 Geräte

Gerät	Typenbezeichnung und Hersteller
Brutschrank	Fa. Heraeus und KB 53 (Binder)
Dialysemembranen	Porengröße 0.025 µm, Durchmesser 25 mm (Millipore)
Digitalkamera	Power Shot A70 (Canon)
Dokumentationssystem für Agarosegele	Computer E.A.S.Y. und Kamera (beide UVP), UV-Schirm (Herolab), Thermodrucker UP-D 895 (SONY)
Elektroporator	Micropulser® (Bio-Rad)
Einmalküvetten	Plastibrand® PS (Brand)
FPLC-Säulen	XK 16 und K 9/15 (Amersham Pharmacia Biotech)
FPLC-System	Pumpen P-500, Gradientenprogrammierer GP-250, Fraktionskollektor Frac-100, Schreiber LKB Rec-102 und 2238 UVCord S II (Amersham Pharmacia Biotech)
FRENCH-Press	FRENCH®-Pressure Cell-Version 5.1 und Standard Cell 40KP (Sim Aminco)
Apparaturen für die Agarose-Gelelektrophorese	Feinmechanik, Philipps-Universität Marburg
Apparaturen für die SDS-PAGE	Mighty Small (Hoefer Scientific)
Elektroporationsküvette	Electroporation Cuvette, 2 mm gap (Molecular Bio Products)
Fluidizer	EmulsiFlex-C5 (Avestin, Ottawa, Kanada)
Gelschüttler	Modell # 3013 (GFL)
Heizblock	BT3 (GRANT Instruments)
HPLC-Säule	C18 Nucleodur-Reversed Phase (Macherey-Nagel)
HPLC-System	Agilent 1100 Series (Hewlett Packard)
Konzentratoren	Vivaspin Concentrator, 6 ml und 20 ml (Vivascience)
Kristallisationszubehör	VDX- / Cryschem-24 well-Platten, Deckgläser (Ø=22 mm) (Hampton Research); Innovaplate®, 96 well (Innovadyne)
Lyophilisationsgerät	GT2 (Leybold-Heraeus), Pumpe AMEB 90 SY4 R3 (AEG)

## 2 Material

Gerät	Typenbezeichnung und Hersteller
MALDI	Biflex (Bruker) mit einer Vakuumanlage der Fa. Pfeiffer
Mikroskope	Z 8 (Leica), Durchleuchtetisch SZ ST und optischer Aufbau SZ 60 (Olympus)
MS-System	Q Star Pulsar i (Applied Biosystems)
pH-Meter	pH 525 (WTW)
Photometer	U-2000 (Hitachi), Ultrospec® 3100 <i>pro</i> (GE Healthcare)
Quarzküvetten	105.200-QS (Hellma)
PCR-Maschine	Robocycler Gradient 96 (Stratagene) und T-personal 48 (Biometra)
Reinstwasser-Anlage	Seralpur Pro90CN (Seral)
Röntgendiffraktometer	Image plate 345 dtb (MAR Research), Drehanode Nonius FR591 (Bruker)
Schüttler	TH 25 (Edmund Bühler), G25 (New Brunswick Scientific)
Spannungsquellen für Agarosegele	Model 453 (ISCO)
Spannungsquellen für SDS-PAGE	EPS 300 (Amersham Pharmacia Biotech)
Spritzenvorsatzfilter	Filtropur S 0.2, Filterporengröße 0.2 µm (Sarstedt)
Ultraschallgerät	Sonoplus HD 200 (Bandelin)
Waagen	PC2200 und Labstyle 54 (Toledo-Mettler)
Wasserbad	NK22 (Haake)
Zentrifugen	Biofuge 15, Biofuge 13R (Heraeus); Centrifuge 5810R mit Rotor A4 62 (Eppendorf); J2-21M/E und J2-HS (Beckman)
Zentrifugenfilter	Zentrifugenfilter bis 0.5 ml, Porengröße 0.2 µm (Millipore)

## 2.2 Chemikalien, Biochemikalien und Enzyme

Für die Herstellung eigener Kristallisationslösungen wurden Chemikalien der Firma FLUKA erworben. Für das Screening zur Identifikation geeigneter Kristallisationsbedingungen wurden Lösungen (*kits*) der Firmen Qiagen und Sigma verwendet, deren Zusammensetzung im Anhang (Kapitel 8.4) aufgelistet sind. Sofern nicht anders beschrieben, wurden sämtliche verwendeten Chemikalien von Sigma-Aldrich in der Reinheitsstufe *pro analysi* bezogen.

Enzyme für molekularbiologische Arbeiten, wie Restriktionsendonukleasen und Ligasen, wurden von New England Biolabs (Ipswich, MA, USA) und DNA-Oligonukleotide von der Firma Invitrogen (Karlsruhe) bezogen. Sequenzierungen von Plasmid-DNA wurden von GATC (Konstanz) bzw. Qiagen (Hilden) durchgeführt.

Daneben wurden folgende Chemikalien und Reagenzien verwendet: BIO-RAD Protein Assay (BIO-RAD, München); IPTG (Gerbü, Gaiberg); Ampicillin und Kanamycin (Applichem, Darmstadt); Bacto-Trypton und Bacto-Hefeextrakt (Difco, Detroit, USA); peqGOLD Protein-Marker I und peqGOLD Range Mix DNA-Leiter (peQLab, Erlangen); Phusion Polymerase (Finnzymes, Espoo, Finnland); L-(+)-Selenomethionin (ACROS Organics, New Jersey, USA).

### 2.2.1 Puffer

#### Auftragspuffer für die Ni<sup>2+</sup>-NTA-Affinitätschromatographie

Menge	Substanz
2.92 g	NaCl (50 mM)
5.96 g	HEPES (25 mM)
dest H <sub>2</sub> O ad 1 L, pH 8, sterilfiltriert	

#### Elutionspuffer für die Ni<sup>2+</sup>-NTA-Affinitätschromatographie

Menge	Substanz
2.92 g	NaCl (50 mM)
5.96 g	HEPES (25 mM)
34.05 g	Imidazol (500 mM)
dest H <sub>2</sub> O ad 1 L, pH 8, sterilfiltriert	

### Puffer für die Gelfiltration

Menge	Substanz
5.84 g	NaCl (100 mM)
2.36 g	HEPES (10 mM)
dest H <sub>2</sub> O ad 1 L, pH 7.5, sterilfiltriert	

Bei den Gelfiltrationsläufen mit Selenomethionin-substituiertem Protein war dem Puffer zusätzlich 1.5 mM DTT zugesetzt worden, um einer Oxidation der Selenatome vorzubeugen.

## 2.3 Medien

### 2.3.1 Luria-Bertani-Medium (LB-Medium)

Menge	Substanz
10 g	Trypton
5 g	Hefeextrakt
10 g	NaCl

Pro Liter anzusetzendem Medium wurden die genannten Mengen in entionisiertem Wasser aufgelöst und auf 1 L aufgefüllt, dann 400 µL 10 M Natronlauge hinzugefügt und die Lösung anschließend für 30 min bei 121 °C und 1.5 bar autoklaviert. Vor dem Animpfen wurden gegebenenfalls benötigte Selektionsmarker zugefügt.

### 2.3.2 LB-Agar-Festmedium

Zur Herstellung von LB-Agar-Festmedium wurde dem LB-Medium vor dem Autoklavieren Agar in einer Endkonzentration von 15 g/L hinzugefügt. Die Zugabe von Selektionsmarkern erfolgte nach Abkühlen des Festmediums auf etwa 50 °C.

### 2.3.3 M9-Minimalmedium

Das M9-Minimalmedium setzt sich folgendermaßen zusammen:

Menge	Substanz
100 mL	M9-Minimalmedium-Stammlösung (autoklaviert)
20 mL	D-(+)-Glucose (400 g/L; sterilfiltriert)
1 mL	MgSO <sub>4</sub> (100 g/L; autoklaviert)
400 µL	CaCl <sub>2</sub> (100 g/L; autoklaviert)
1 mL	Biotin/Thiamin (je 2 g/L; sterilfiltriert)
888 mL	bidest. Wasser (autoklaviert)

Die M9-Minimalmedium-Stammlösung enthält (pro Liter):

Menge	Substanz
105 g	K <sub>2</sub> HPO <sub>4</sub> * 3 H <sub>2</sub> O
45 g	KH <sub>2</sub> PO <sub>4</sub>
5 g	Na <sub>3</sub> -Citrat
10 g	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>

Vor dem Animpfen mit einer Starterkultur wurden die gegebenenfalls benötigten Selektionsmarker zugefügt. Des Weiteren wurden zur Suppression der zelleigenen Methionin-Biosynthese zwei Lösungen von Aminosäuren in bidest. Wasser benötigt:

Aminosäure-Mix I	L-Lys, L-Phe, L-Thr	(je 10 g/L; sterilfiltriert)
Aminosäure-Mix II	L-Ile, L-Leu, L-Val	(je 5 g/L; sterilfiltriert)

Für die Synthese von Selenomethionin-markiertem Protein wurde diese Aminosäure in Form einer wässrigen, sterilfiltrierten Lösung (5 g/L) zugegeben.

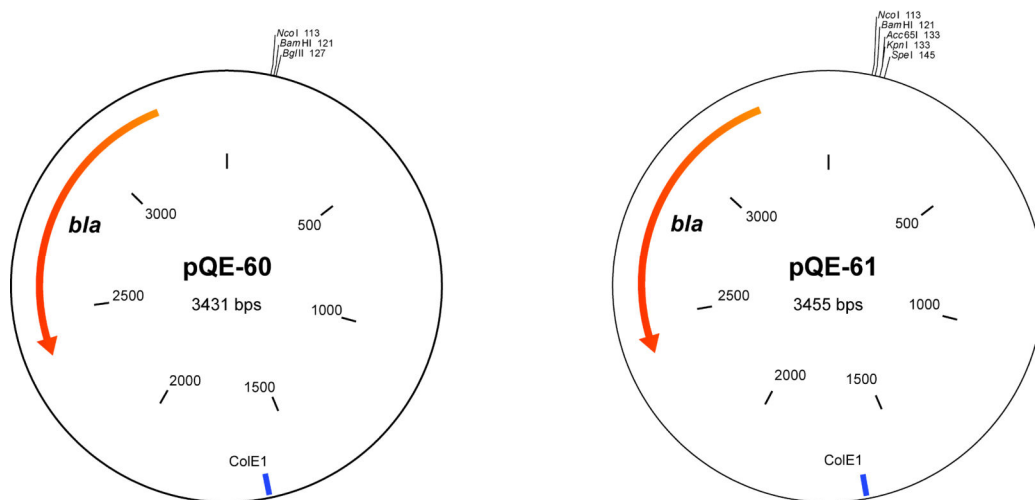
## 2.4 Vektoren

### pQE-60

Dieser Vektor der Firma Qiagen eignet sich für die Überproduktion rekombinanter Proteine in *E. coli*. Er verfügt über eine *multiple cloning site* (MCS) mit drei Restriktionsschnittstellen. Außerdem kodiert er für einen C-terminalen His<sub>6</sub>-tag, welcher die anschließende Reinigung des rekombinanten Proteins mittels Ni<sup>2+</sup>-NTA-Affinitätschromatographie ermöglicht. Das  $\beta$ -Lactamase-Gen (*bla*) verleiht Resistenz gegenüber Ampicillin (Arbeitskonz. 100  $\mu$ g/mL) und gestattet somit die Selektion transformierter Zellen.

### pQE-61

Hierbei handelt es sich um eine Variante des oben beschriebenen, kommerziell erhältlichen pQE-60. Die beiden Vektoren unterscheiden sich in ihrer MCS. Im pQE-61 ist die *Bgl*II-Schnittstelle durch ein 24 bp-Insert unterbrochen, welches wiederum Schnittstellen für die Restriktionsenzyme *Acc*65I, *Kpn*I sowie *Spe*I enthält.



**Abbildung 2.1: Karten der verwendeten Vektoren, pQE-60 und pQE-61**

Neben den Restriktionsschnittstellen der *multiple cloning sites* sind auch die Positionen des Replikationsursprungs (ColE1) und des  $\beta$ -Lactamase kodierenden Gens angezeigt.



## 2.5 Bakterienstämme

### 2.5.1 Klonierungsstämme

#### DH5 $\alpha$ <sup>®</sup>

Dieser *E. coli*-Stamm wird von Invitrogen vertrieben und wurde für Klonierungsarbeiten verwendet. Der Genotyp dieses Stammes lautet:

*E. coli* K-12 F<sup>-</sup>,  $\phi$ 80*lacZ* $\Delta$ M15,  $\Delta$ (*lacZYA-argF*), U169, *recA1*, *endA1*, *hsdR17*,  
(r<sub>k</sub><sup>-</sup>,m<sub>k</sub><sup>+</sup>), *phoA*, *supE44*  $\lambda^-$ , *thi1*, *gyrA96*, *relA1*

### 2.5.2 Expressionsstämme

#### M15 [pREP4]

Sämtliche Überexpressionen dieser Arbeit wurden heterolog in *E. coli* M15 [pREP4] durchgeführt. Bei diesem Stamm handelt es sich um einen von der Firma Qiagen vertriebenen zu pQE-Vektoren kompatiblen *E. coli* K-12-Expressionsstamm, der den folgenden Phänotyp aufweist:

*E. coli* K-12 NaI<sup>S</sup>, Str<sup>S</sup>, Rif<sup>S</sup>, Thi<sup>-</sup>, Lac<sup>-</sup>, Ara<sup>+</sup>, Gal<sup>+</sup>, Mtl<sup>-</sup>, F<sup>-</sup>, RecA<sup>+</sup>, Uvr<sup>+</sup>, Lon<sup>+</sup>

Um eine unerwünschte, vorzeitige Genexpression während der frühen Wachstumsphase zu unterbinden, wird die Expression durch das Operator-System des *lac*-Operons supprimiert. Das *lacI*-Gen ist im Genom des Expressionsstammes M15 nicht kodiert. Dieses Gen ist jedoch im Helferplasmid pREP4 kodiert, welches in den Zellen vorliegt. Das Helferplasmid verleiht darüber hinaus eine Resistenz gegenüber Kanamycin (Arbeitskonzentration 25  $\mu$ g/mL).



## 3 Methoden

### 3.1 Molekularbiologische Methoden

#### 3.1.1 Synthese von DNA-Oligonukleotiden

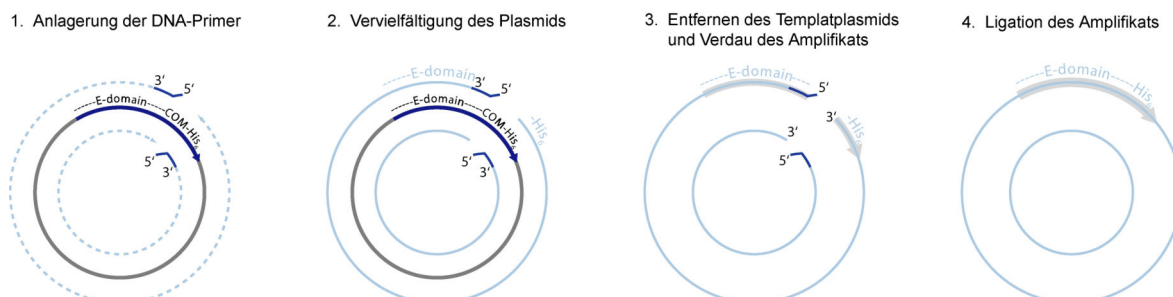
Die Sequenz der DNA-Primer wurde anhand der Templat-DNA unter Berücksichtigung der *nearest neighbour*-Methode (<http://www.idtdna.com/>) bestimmt. Dabei wurde darauf geachtet, dass sich die berechneten Schmelztemperaturen der beiden Primer in beiden Hybridisierungsschritten, d. h. sowohl der Anlagerung der Primer an Templat-DNA einerseits (nur homologe Sequenzbereiche) als auch den nachfolgenden Anlagerungsschritten an das Amplifikat andererseits (über die gesamte Primerlänge), um weniger als 2 °C unterscheiden.

#### 3.1.2 Polymerase-Kettenreaktion (PCR)

Die Polymerase-Kettenreaktion ermöglicht die Amplifikation ausgewählter DNA-Sequenzen [Mullis, 1987]. Durch wiederholtes Denaturieren der DNA-Doppelstränge können sich bei jedem Abkühlen die DNA-Primer an das Templat anlagern. Anschließend werden die gebundenen DNA-Primer durch die Polymerase verlängert. Die Sequenz der aufeinander zu gerichteten Primer bestimmt die Position ihrer Anlagerung und somit auch die zu amplifizierende Sequenz.

#### 3.1.3 Inverse Polymerase-Kettenreaktion (iPCR)

Die inverse PCR (siehe Abbildung 3.1) beschreibt eine Variante der PCR, bei der die Primer voneinander weg gerichtet sind. Wie unten dargestellt, wird nach Anlagerung der Primer die Sequenz des Plasmids zwischen den Primern amplifiziert. Der kurze Sequenzabschnitt zwischen



**Abbildung 3.1: Schema der inversen PCR zur Manipulation einer Sequenz im Vektor**

Als Beispiel dient die in dieser Arbeit durchgeführte Entfernung des COM-Sequenzbereichs. Das ursprüngliche Templat und das Produkt sind dunkelgrau bzw. hellblau dargestellt. Der das rekombinante Protein kodierende Bereich ist als dunkelblauer Pfeil hervorgehoben.



verkürzte Inkubationszeit der alkalischen Lyse. Sobald nach wiederholtem Invertieren der Überstand sichtbar klarer wurde, wurde die Lyse durch Zugabe von Puffer P3 gestoppt. Die zweite Änderung betrifft das Volumen des Puffer EB zur Elution des Plasmids. Um eine höhere DNA-Konzentration zu erreichen, wurde DNA nur mit 30 µL Puffer EB eluiert.

### **3.1.6 Hydrolytische Spaltung von DNA**

Für die sequenzspezifische, hydrolytische Spaltung von DNA-Proben wurden Typ II Restriktionsendonukleasen verwendet. Dazu wurde die Probe mit den entsprechenden Enzymen und dem für diese Enzyme vom Hersteller empfohlenen Reaktionspuffer für 2 Stunden inkubiert. Anschließend wurden die Enzyme inaktiviert und die Reaktionsansätze per Agarosegelelektrophorese analysiert.

### **3.1.7 Ligation**

Zueinander komplementäre, hydrolytisch gespaltene DNA-Fragmente können zu Duplex-DNA hybridisieren. In dieser Duplex-DNA sind jedoch die Phosphodiesterbindungen im DNA-Rückgrat noch nicht ausgebildet. Die kovalenten Bindungen zwischen den 5'-Phosphat- und den 3'-Hydroxygruppen können in einer Enzym-katalysierten, ATP-abhängigen Reaktion ausgebildet werden. Für diese als Ligation bezeichnete Reaktion wurde die T4-DNA-Ligase als Enzym eingesetzt.

### **3.1.8 Entsalzung von Ligationsansätzen**

Vor der Durchführung elektrochemischer Transformationen mussten die zuvor durchgeführten Ligationsansätze entsalzt werden. Dazu wurden die Proben auf Dialysemembranen eine Stunde gegen destilliertes Wasser dialysiert.

### **3.1.9 Agarosegelelektrophorese**

Die analytische Elektrophorese zur Auftrennung von DNA-Fragmenten wurde mit 1%igen Agarosegelen unter Verwendung eines TBE-Puffersystems und mit einer Stromstärke von 100 mA durchgeführt [Sambrook, 1989]. Die DNA-Proben wurden zuvor mit dem fünffachen Volumen eines wässrigen 6x DNA-Probenpuffers versetzt, dessen Zusammensetzung nachfolgend aufgeführt ist. Die Gele wurden in einer wässrigen Ethidiumbromid-Lösung (200 µg Ethidium-

bromid/200 mL dest. Wasser) gefärbt. Anschließend wurde der Hintergrund in Wasser entfärbt. Die eingefärbten DNA-Banden wurden unter UV-Licht sichtbar gemacht und photographiert.

Menge	Substanz
0.25 %	Xylenxanol FF
15 %	Ficoll (Typ 400)

#### 3.1.10 Photometrische DNA-Konzentrations- und Reinheitsbestimmung

Die Konzentrationsbestimmung von doppelsträngiger Plasmid-DNA wurde per Photometer durchgeführt. Dazu wurden in Quarzküvetten 2 µL der zu bestimmenden DNA-Probe mit 198 µL destilliertem Wasser verdünnt. Nach Messung der Extinktion bei einer Wellenlänge von 260 nm ( $E_{260}$ ) ließ sich die DNA-Konzentration ( $c$ ) mit Hilfe der nachfolgenden Formel berechnen [Mülhardt, 2006].

$$c \text{ [}\mu\text{g/mL]} = E_{260} \cdot 5000$$

Des Weiteren ließ sich mit Hilfe des Quotienten der Extinktionen bei 260 nm und 280 nm,  $E_{260}/E_{280}$ , eine Aussage über die Reinheit der Probe treffen.

#### 3.1.11 DNA-Sequenzanalyse

Sämtliche DNA-Sequenzanalysen wurden nach der Kettenabbruchmethode mit basenspezifischen Fluoreszenzmarkern von den Firmen GATC und Qiagen durchgeführt. Anhand der erhaltenen Interferogramme wurde die Sequenz der Probe auf Richtigkeit überprüft.

#### 3.1.12 Darstellung und Transformation kompetenter *E. coli*-Zellen

Chemisch kompetente bzw. elektrokompente *E. coli*-Zellen wurden gemäß der Vorschrift von Sambrook *et al.* hergestellt [Sambrook, 1989], unter Kühlung in Proben zu je 50 µL aliquotiert und in flüssigem Stickstoff schockgefroren. Die tiefgefrorenen Zellen wurden bei -80 °C gelagert. Zur Verwendung wurden sie kurz vor der Verwendung in Eiswasser aufgetaut.

### 3.1.13 Stammhaltung

Zur langfristigen Lagerung der Zellen wurden Glycerinstocks angesetzt. Dazu wurde ein Teil einer Übernachtskultur mit dem doppelten Volumen 87%igen, sterilen Glycerins versetzt, gründlich durchmischt und anschließend in flüssigem Stickstoff schockgefroren. Die Proben wurden bei -80 °C gelagert. Daneben wurde von sämtlichen kristallisierten Proteinen eine Plasmidpräparation angefertigt und diese bei -20 °C gelagert.

## 3.2 Proteinchemische Methoden

### 3.2.1 Heterologe Genexpression in *E. coli*

Für die meisten Arbeiten mit *E. coli*-Kulturen wurde Luria-Bertani-Medium (LB-Medium) verwendet. Ausgehend von einer hochgewachsenen Starterkultur wurde vorgewärmtes LB-Medium (mit den ggf. erforderlichen Selektionsmarkern) in einem Verhältnis von 1:100 angeimpft und im Schüttelinkubator herangezogen (30 °C, 225 rpm). Die Synthese des rekombinanten Proteins wurde bei einer  $OD_{595}=0.6$  durch Zugabe von IPTG induziert und die Kultur für weitere 3 Stunden inkubiert (30 °C, 225 rpm).

Daneben fand M9-Minimalmedium Verwendung, wenn für röntgenkristallographische Zwecke (experimentelle Phasierung mittels MAD) Selenomethionin-markiertes Protein hergestellt werden sollte. Um, wie von Van Duyne *et al.* beschrieben, die zelleigene Methionin-Biosynthese zu unterdrücken, wurden vor der Induktion mit IPTG die unten beschriebenen, sterilfiltrierten Aminosäurenlösungen (je 10 mL/L Kultur) und später die Selenomethionin-Lösung zugegeben [Van Duyne, 1993].

Wie schließlich zuerst von Drake *et al.* [Drake, 2006] publiziert, wurde beobachtet, dass die Expression PCP-Domänen-enthaltender rekombinanter Proteine in eisen-defizientem Medium (z. B. M9-Minimalmedium) deren *holo*-Form liefert, da unter diesen Bedingungen der Wirtsorganismus *E. coli* mit Hilfe der zelleigenen nichtribosomalen Synthese das Siderophor Enterobactin synthetisiert. Die mit der Enterobactin-Biosynthese assoziierte 4'-Phosphopantetheinyl-Transferase EntD modifiziert ebenfalls die PCP-Domäne des rekombinanten Proteins.

Es ließ sich massenspektrometrisch zeigen, dass Zugabe von 20 mg/L  $FeSO_4$  die durch Eisenmangel bedingte Phosphopantetheinylierung der rekombinanten PCP-Domäne inhibiert. Die Enterobactin-Synthetasen und damit assoziierte Enzyme werden dann nicht exprimiert, wodurch das rekombinante Protein in der *apo*-Form, also unmodifiziert, erhalten wird.

### 3.2.2 Präparation von *E. coli* Rohzelleextrakten

Für analytische Zwecke wurde das Zellpellet einer 50 mL Kultur in 500  $\mu$ L Auftragspuffer aufgenommen, mit 5  $\mu$ L 0.1 M EDTA und 1  $\mu$ L 50 mg/mL Lysozym versetzt und unter gelegentlichem Mischen 15 min auf Eis inkubiert. Anschließend wurden die Zellen unter ständiger Eiskühlung mittels Ultraschall aufgeschlossen (5 Zyklen à 20 Pulse, unterbrochen durch Pausen à 30 s; Bandelin Sonoplus HD 200; Einstellungen: *Program 72/D* und *cycle 40*).

Im präparativen Maßstab wurden in Puffer resuspendierte und schockgefrorene Zellen zunächst auf Eis aufgetaut. Der präparative Zellaufschluss erfolgte entweder an einer French Press® oder an einem Fluidizer.

### 3.2.3 Affinitätschromatographie

Alle rekombinant produzierten Proteine enthielten an ihrem C-Terminus sechs zusätzliche, durch den Vektor kodierte Histidinreste (*His<sub>6</sub>-tag*). Dies ermöglichte die effiziente Reinigung mittels  $\text{Ni}^{2+}$ -Nitrilotriessigsäure-Affinitätschromatographie ( $\text{Ni}^{2+}$ -NTA; allgemein: IMAC, *immobilized metal ion affinity chromatography*), da die Histidinreste unter den Auftragsbedingungen an die  $\text{Ni}^{2+}$ -Ionen koordinieren. Unter den Elutionsbedingungen wird das Protein durch Imidazol von den Bindungsstellen der Säulenmatrix verdrängt.

Die Reinigungen wurden stets bei 4-8 °C durchgeführt. Mit Hilfe des FPLC-Systems wurde das Protein nach Auftragen (ohne Imidazol) und Waschen (15 mM Imidazol) durch Anlegen eines Imidazolgradienten eluiert (0-500 mM Imidazol über 60 mL; Säulenvolumen 8 mL).

Bei den Reinigungen von PCP-C Bidomänen wurde das Puffersystem bei pH 7 verwendet, für die späteren Reinigungen der Epimerisierungsdomänen waren die Puffer auf pH 8 eingestellt.

### 3.2.4 Ultrafiltration

Diese Methode wurde zum Aufkonzentrieren von Proteinlösungen verwendet. An einer Membran, die aufgrund ihrer definierten Porengröße für die Lösungsmittelmoleküle, nicht aber für das zu konzentrierende Protein durchlässig ist, wird das Lösungsmittel selektiv entfernt. Infolge dessen steigt im Retentat die Proteinkonzentration.

Für das Aufkonzentrieren wurden Amicon Ultra Konzentratoren (Millipore) in einer Kühlzentrifuge (Eppendorf 5417 C) bei 4 °C verwendet. Je nach Molekulargewicht der Proteinprobe kamen Konzentratoren mit Ausschlussgrößen von 10 kDa, 30 kDa oder 50 kDa zum Einsatz. Für sehr kleine Probenvolumina wurden auch Amicon Microcon Konzentratoren (Millipore) eingesetzt.



### 3.2.5 Gelfiltration

Die Gelfiltration wurde zur Trennung von Proteingemischen unter nativen Bedingungen angewendet. Das Prinzip der Trennung beruht auf der größenabhängigen Wechselwirkung zwischen den Probenmolekülen und der porösen Struktur des Säulenmaterials. Die Ausschlussgröße und der Trennbereich werden durch die Porengröße des verwendeten Säulenmaterials bestimmt. Meist wurde eine Superdex 200-Säule (Säulenkörper XK16, 120 mL Gelbettvolumen) verwendet und diese bei 8 °C an einem FPLC-System mit Flussraten von 0.5-1 mL/min betrieben.

Nach Kalibration der Säule mit Proteinen des *low molecular weight calibration kits* und des *high molecular weight calibration kits* (Amersham) war es möglich, die apparente Proteinmasse abzuschätzen. Daraus ließ sich dann auch der Oligomerisierungszustand des Proteins ableiten.

### 3.2.6 Entsalzung und Umpufferung

Seltener wurde für kleine Mengen eine PD-10 Säule (GE Healthcare) nach den Angaben des Herstellers verwendet. Hierbei handelt es sich eine kleine Gelfiltrationssäule, die eine brauchbare Trennung zwischen Makromolekülen und niedermolekularen Substanzen liefert und sich deswegen beispielsweise für Pufferwechsel eignet.

### 3.2.7 Sterilfiltration

Verschiedene, temperaturempfindliche Lösungen für Arbeiten mit *E. coli*-Kulturen, wie beispielsweise Stammlösungen von Antibiotika und des Induktors IPTG, Komponenten des M9-Minimalmediums, sowie Puffer und Proteinlösungen der Proteinaufreinigung wurden sterilfiltriert. Des Weiteren wurden auch die bei der Kristallisation verwendeten Puffer und Proteinlösungen sterilfiltriert, um feste Bestandteile und agglomeriertes Protein zu entfernen, die als unerwünschte Kristallisationskeime fungieren könnten. Bei größeren Volumina wurde mit einer Filterapparatur der Firma Millipore gearbeitet. Bei kleineren Mengen kamen Spritzenvorsatzfilter und Zentrifugensterilfilter (je 0.22 µm Porengröße) zum Einsatz.

### 3.2.8 Proteinkonzentrationsbestimmung

Die Proteinkonzentration wurde jeweils vor und nach Aufreinigungsschritten, während des Aufkonzentrierens sowie bei Kristallisationsexperimenten nach der Sterilfiltration bestimmt.

Die Konzentration von Proteinlösungen wurde photometrisch nach der Methode von Bradford durchgeführt. Das *BIO-RAD Protein Assay* Konzentrat (BioRad) wurde mit der vierfachen Menge dest. Wasser verdünnt, filtriert und bis zur Verwendung lichtgeschützt bei Raumtemperatur aufbewahrt. In den jeweils zweifach durchgeführten Bestimmungen wurde zu 1 mL dieser Testlösung 1 µL der Proteinlösung pipettiert, gründlich durchmischt und 5 min bei Raumtemperatur inkubiert. Die Messung der Absorption erfolgte bei einer Wellenlänge von 595 nm. Zur Ableitung der Proteinkonzentration aus den gemessenen Absorptionen diente eine Kalibriergerade, die zuvor mit BSA-Lösungen definierter Konzentrationen erstellt wurde.

### 3.2.9 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)

Die Analyse von Proteinproben aus Rohzelleextrakten oder nach Reinigungsschritten erfolgte mittels diskontinuierlicher SDS-PAGE nach der Methode von Lämmli [Laemmli, 1970]. Die Proben wurden mit SDS-Probenpuffer versetzt und vor dem Auftragen 5 min bei 95 °C erhitzt, um eine vollständige Denaturierung zu gewährleisten.

## 3.3 Proteinkristallisation und Strukturbestimmung

Da sich die geeigneten Kristallisationsbedingungen für verschiedene Proteine unterscheiden und die Methoden zur computergestützten Vorhersage unausgereift sind, haben sich in der Praxis *sparse matrix screens* durchgesetzt [Jancarik, 1991]. In Kombination mit der Automatisierung, die sowohl einen hohen Probendurchsatz, als auch Tropfengrößen im Sub-Mikroliter-Maßstab ermöglicht, wird eine große Anzahl Proben mit verschiedenen Bedingungen präpariert. Im Anschluss werden die Ansätze – ggf. über einen längeren Zeitraum – kontrolliert, um mögliche geeignete Kristallisationsbedingungen zu identifizieren.

Prinzipiell wurden die Ansätze bei konstanter Temperatur mit thermisch äquilibrierten Kristallisationslösungen und auf Eis gekühlten Proteinproben präpariert.

### 3.3.1 *hanging drop*- und *sitting drop*-Kristallisation

Hierbei handelt es sich um zwei weitverbreitete Methoden zur Proteinkristallisation. In einem Ansatz liegt ein Tropfen, bestehend aus der Proteinlösung und der Kristallisationslösung, neben einem vergleichsweise großen Reservoir der entsprechenden Kristallisationslösung. Nach Verschließen des Ansatzes bildet sich über die Dampfphase zwischen den beiden flüssigen Phasen ein Gleichgewicht aus. Im Laufe dieser langsamen Gleichgewichtseinstellung verliert der Tropfen meist an Volumen. Dies hat zur Folge, dass die Konzentration der in ihm gelösten Stoffe (Salze, Fällungsmittel, Protein) ansteigen. Unter geeigneten Bedingungen kann dabei die Proteinkonzentration in einen übersättigten Bereich fallen und Kristalle entstehen.

Bei der *hanging-drop* Methode [Davies, 1971] wird der Probentropfen auf ein Deckglas pipettiert. Dieses wird vorsichtig umgedreht und damit die gefettete Öffnung des Probenraums geschlossen, so dass der Tropfen innen herunterhängt. Bei der *sitting-drop* Methode liegt der Probentropfen in einer Mulde neben dem Reservoir vor, und der Probenraum wird durch eine selbstklebende, transparente Folie verschlossen. Wegen der Möglichkeit zur Automatisierung, wurden die anfänglichen Screens in *sitting-drop*-Ansätzen durchgeführt. Hierbei fanden Crystalquick Plus- und Innova-Platten Einsatz, die jeweils mit view seal-Folien abgedichtet wurden. Die manuelle Reproduktion und Optimierung erfolgte in Cryschem-Platten (*sitting-drop*) mit CrystalClear Klebeband oder in VDX-Platten (*hanging-drop*) mit silikonisierten Deckgläsern.

### 3.3.2 *Soaking*-Experimente

*Soaking* beschreibt ein Verfahren zur Derivatisierung eines bestehenden Kristalls. Hierbei werden zum einen Substrate oder Inhibitoren verwendet, um Informationen über deren Assoziation mit dem Protein zu erhalten. Zum anderen werden auch Schwermetall-Verbindungen bzw. -Ionen eingesetzt, die zur experimentellen Phasenbestimmung durch (multiplen) isomorphen Ersatz oder (multiple) anomale Dispersion genutzt werden können. In der Praxis wird dabei ein Kristall aus seiner Mutterlösung in einen Tropfen vergleichbarer Zusammensetzung überführt, der zusätzlich die zu inkorporierende Verbindung enthält. Abhängig von der Größe der Solvenskanäle können die gelösten Verbindungen in den Kristall diffundieren, wo sie (Zugänglichkeit der Bindungsstelle vorausgesetzt) an das Protein binden können.

### 3.3.3 Seeding-Experimente

Unter *Seeding* versteht man das Animpfen eines neuen Kristallisationsansatzes mit Impfkristallen. Man unterscheidet dabei zwischen dem *Makro-Seeding*, bei dem intakte Proteinkristalle transferiert werden, um größer zu werden, und das *Mikro-Seeding*, bei dem Fragmente von zuvor zerkleinerten Proteinkristallen in den Ansatz eingebracht werden. Diese sollen als Kristallisationskeime fungieren und so größere oder gleichmäßigere gewachsene Kristalle liefern.

Beide Techniken wurden mit den Proteinen PCP-C und TycB3-E durchgeführt. Für das *Mikro-Seeding* wurden die vorhandenen Kristalle entweder zwischen zwei Deckgläsern oder mit Hilfe von *Mikrotools* im Kristallisationstropfen selbst zerdrückt. Der Transfer der Impfkristall-Suspension erfolgte vorzugsweise mit einem Katzenhaar. Im Falle der TycB3-E-Domäne, die nur in *sitting drop*-Ansätzen kristallisierte, wurde wegen der erschwerten Zugänglichkeit des Tropfens in der Tropfenaufgabe alternativ ein *Mikrotool* verwendet.

Darüberhinaus wurde beim *Mikro-Seeding* auch der Einfluss der Konzentration der verwendeten Impfkristall-Suspension untersucht. Zum einen wurden unterschiedliche Verdünnungen eingesetzt, zum anderen wurden mehrere Tropfen unmittelbar nacheinander mit einem Katzenhaar angeimpft, so dass die Anzahl der eingebrachten Impfkristalle abnahm.

### 3.3.4 Testen von Cryo-Bedingungen

Um die prinzipielle Eignung eines Cryopuffers für cryokristallographische Zwecke, d. h. ausreichende Unterdrückung der Ausbildung von Eisstrukturen, zu untersuchen, wurde ein Nylon-Loop mit dem Cryopuffer benetzt und im Cryo-Gasstrom bei 100 K schockgefroren. Nach Aufnahme eines Diffraktionsbildes wurde anhand des Auftretens bzw. des Nicht-Auftretens der für Eis charakteristischen Ringe die Eignung des Puffers beurteilt.

Des Weiteren mussten die Kristalle die Wechsel in veränderten Bedingungen des Cryopuffers ertragen. Für *Soaking*-Experimente mit Substraten, aus denen die Kristalle direkt schockgefroren wurden, wurde das *Soaking* im Cryopuffer durchgeführt. Um den Einfluss der längeren Inkubation in Cryopuffer einzuschätzen, wurden Kristalle über Nacht in Cryopuffer inkubiert und anschließend am Diffraktometer getestet.

### 3.3.5 Datensammlung und Datenreduktion

Die Cryoschleifen wurden auf dem Goniometerkopf angebracht. Nach Zentrierung der Kristalle wurden einige Diffraktionsbilder aufgenommen. Aus diesen ließen sich die Zellparameter, die Raumgruppe und die zu erwartende Auflösung abschätzen. Die Indizierung wurde meist mit dem

Programm MOSFLM [Leslie, 1992] durchgeführt. Alternativ wurde am Synchrotron des ESRF (Grenoble, Frankreich) mit der dort installierten Software DNA (<http://www.dna.ac.uk/index.html>) gearbeitet. Beide Programme schlugen auch Strategien für die Datensammlung vor. Die aufgenommenen Datensätze wurden mit den Programmen MOSFLM oder XDS [Kabsch, 1993] prozessiert. Für die anschließende Skalierung und Datenreduktion wurde das Programm SCALA [CCP4, 1994] verwendet.

### 3.3.6 Suche nach Sequenz- und Strukturhomologen

Sequenzhomologe wurden mit Hilfe des BLAST-Servers ermittelt. Darüber hinaus wurden Strukturhomologe durch die VAST-Suche [Gibrat, 1996] in der MMDB [Wang, 2007] ermittelt. Beide Suchfunktionen sind über die Website des National Center of Biotechnology Information verfügbar (<http://www.ncbi.nih.gov/blast/>).

### 3.3.7 Strukturlösung mittels *molecular replacement*

Für die Durchführung des *molecular replacement* benötigt man eine Struktur, die als Modell der zu lösenden Struktur dient. Die Platzierung des Suchmodells in der Einheitszelle ist durch sechs Parameter bestimmt, drei Rotationsparameter und drei Translationsparameter. Da eine einfache Suche nach diesen sechs Parametern sehr rechenintensiv ist, führen die *molecular replacement*-Programme sequentiell eine Rotationsfunktion und anschließend eine Translationsfunktion durch, was die Anzahl der erforderlichen Rechenoperationen deutlich begrenzt. Die Bewertung der Suchmodellorientierungen erfolgt durch Vergleich der entsprechenden Pattersonfunktion des Suchmodells mit der des zu lösenden Datensatzes. Bei erfolgreichem Ausgang der Rechnungen wird eine Struktur erhalten, die ggf. noch verfeinert werden muss.

Der Erfolg eines Strukturlösungsversuchs per *molecular replacement* hängt meist von der Qualität des verwendeten Strukturmodells ab. Ein gutes Modell ist der zu lösenden Struktur strukturell ähnlich. Zur Abschätzung der Brauchbarkeit wird häufig der Grad der Sequenzidentität bzw. -ähnlichkeit herangezogen. Letztendlich ist ein Modell dann gut, wenn es zur Strukturlösung führt.

### 3.3.8 Strukturlösung mittels multipler, anomaler Dispersion (MAD)

In kristallographischen Experimenten zeigen Schweratome, wie beispielsweise Selenatome, im Gegensatz zu den Atomen leichter Elemente, wie beispielsweise Kohlenstoff, Sauerstoff und Stickstoff, einen Effekt, der als anomale Dispersion bezeichnet wird. Dieser tritt auf, wenn die

Energie der verwendeten Röntgenstrahlung in der Nähe einer Absorptionskante des untersuchten Elements liegt. Der insgesamt schwache Effekt zeigt eine empfindliche Abhängigkeit von der Wellenlänge (=Energie) der Röntgenstrahlung und durch Messung bei verschiedenen Wellenlängen lassen sich die Positionen der Schweratome experimentell bestimmen. Meist werden Datensätze bei drei Wellenlängen aufgenommen. Man spricht dann von multipler, anomaler Dispersion (MAD). Da schließlich auch die Phasen der Strukturfaktoren näherungsweise ermittelt werden können, erlaubt diese Methode die direkte Lösung des Phasenproblems.

#### **3.3.9 Modellbau und Strukturverfeinerung**

Die anfänglichen Strukturmodelle wurden stets mit dem Programm ARP/wARP [Perrakis, 2001] erhalten. Unter Verwendung des Programms Coot [Emsley, 2004] wurde das jeweilige Strukturmodell kontrolliert und ggf. manuell verbessert. Daneben wurden auch Wassermoleküle und Ionen in die Elektronendichtekarten eingepasst. Die Verfeinerung erfolgte mit Refmac5 [Murshudov, 1997].

#### **3.3.10 Moleküldynamische Simulation**

Die moleküldynamischen Simulationen der Fengycin Thioesterasedomäne wurden in Zusammenarbeit mit Herrn Prof. Dr. Essen mit dem Programm NAMD durchgeführt [Phillips, 2005]. Für die anschließende Analyse und sowie die graphische Darstellung der erhaltenen Trajektorien wurde das Programm VMD verwendet [Humphrey, 1996].

#### **3.3.11 Berechnung von Protonierungszuständen**

Um die Protonierungszustände der funktionellen Gruppen saurer und basischer Seitenketten abzuschätzen, wurde der H++-Server (<http://biophysics.cs.vt.edu/H++>) genutzt [Gordon, 2005]. Das automatisierte Programm entfernt zunächst sämtliche kristallographisch beobachteten Wassermoleküle und Heteroatome aus der pdb-Datei. Basierend auf der Kristallstruktur werden dann nach der *continuum solvent*-Methode [Bashford, 1990] und dem Poisson-Boltzmann-Modell die pK<sub>a</sub>-Werte aller titrierbaren Gruppen und somit deren Protonierung bei einem wählbaren pH-Wert berechnet. Bei der Durchführung wurden als Parameter ein pH-Wert von 7 und eine Salzkonzentration von 0.15 M sowie eine externe bzw. interne Dielektrizitätskonstante von 80 bzw. 6 gewählt.

### **3.3.12 Graphische Darstellungen**

Sämtliche Darstellungen von Proteinstrukturen wurden mit dem Programm Pymol 0.97 [DeLano, 2002] angefertigt. Zur Bearbeitung von Bildern und Graphiken wurde mit den Programmen ISIS Draw, Adobe Photoshop 7.0 und Adobe Illustrator CS gearbeitet.





## 4 Ergebnisse

Die vorliegende Arbeit beschäftigt sich mit der röntgenkristallographischen Aufklärung der Domänenstrukturen nichtribosomaler Peptidsynthetasen. Dazu wurde eine Vielzahl rekombinanter Proteine exprimiert, aufgereinigt und auf ihre Eignung für kristallographische Untersuchungen hin getestet. Die bearbeiteten, rekombinanten Proteine und die jeweils durchgeführten Arbeitsschritte sind in der folgenden Tabelle dargestellt.

**Tabelle 1: Übersicht über die bearbeiteten rekombinanten Proteine nichtribosomaler Peptidsynthetasen**

Protein, Synthetase, Organismus	Vektor, Länge, Masse	Expression	Aufreinigung	Kristallisation	Optimierung	Diffraktion	pdb entry (Auflösung)
<b>Thioesterasedomänen</b>							
<b>Fen TE</b> Fengycin Synthetase B <i>B. subtilis</i> F29-3	pQE-61 244 aa 27.5 kDa	✓	✓	✓ <sup>*1</sup>	✓	✓	2cb9 (nativ, 1.8 Å) 2cbg (PMS-Komplex, 2.5 Å)
<b>Bac PCP-TE</b> Bacitracin Synthetase B <i>B. licheniformis</i> ATCC 10716	pQE-60 327 aa 37.6 kDa	✓	✓	✗	—	—	—
<b>CDA TE</b> CDA Synthetase III <i>S. coelicolor</i> A3(2)	pQTEV	✓ <sup>*2</sup>	✓	✗	—	—	—
<b>Kondensationsdomänen</b>							
<b>TycC6 C</b> Tyrocidin Synthetase C <i>B. brevis</i> ATCC 8185	pQE-61 458 aa 51.6 kDa	✓	✓	✗	—	—	—
<b>TycC6 PCP-C</b> Tyrocidin Synthetase C <i>B. brevis</i> ATCC 8185	pQE-61 540 aa 61.0 kDa	✓	✓	✓	✓	✓	2jgp (nativ, 1.85 Å)
<b>TycC6 PCP-C (H224A mutant)</b> Tyrocidin Synthetase C <i>B. brevis</i> ATCC 8185	pQE-61 540 aa 60.9 kDa	✓	✓	✓	✓	✓	nativ, 4 Å
<b>TycC6 PCP-C-A-PCP</b> Tyrocidin Synthetase C <i>B. brevis</i> ATCC 8185	pTrc His2 TOPO 1150 aa 127.8 kDa	✓	✓	✗ <sup>*3</sup>	—	—	—
<b>Zyklisierungsdomänen</b>							
<b>BacA1-2 A-PCP-Cy</b> Bacitracin Synthetase A <i>B. licheniformis</i> ATCC 10716	pQE-60 1066 aa 121.2 kDa	✓	✓	✗	—	—	—
<b>BacA1-2 PCP-Cy</b> Bacitracin Synthetase A <i>B. licheniformis</i> ATCC 10716	pQE-60 530 aa 61.8 kDa	✓	✓	✗	—	—	—
<b>BacA2 Cy</b> Bacitracin Synthetase A <i>B. licheniformis</i> ATCC 10716	pQE-60 452 aa 52.9 kDa	✓	✓	✗	—	—	—
<b>BacA2 Cy-A-PCP</b> Bacitracin Synthetase A <i>B. licheniformis</i> ATCC 10716	pQE-60 1053 aa 120.4 kDa	✓	✓	✗	—	—	—
<b>Adenylierungsdomänen</b>							
<b>MtaD A(Ox)</b> Myxothiazol Synthetase D <i>Stigmatella aurantiaca</i> DW4/3-1	pET-20b(+) 817 aa 88.9 kDa	✓	✓	✗	—	—	—

### Anmerkungen:

- ★1 Die Größe der erhaltenen Kristalle ist stark abhängig von der Probenzusammensetzung hinsichtlich des Oligomerisierungszustands des Proteins FenTE.
- ★2 Expression durchgeführt von Dr. Jan Grünewald (Arbeitsgruppe Prof. Marahiel).
- ★3 Kleine Kristalle wurden mit *holo*-Protein PCP-C-A-PCP beobachtet.

Fortsetzung siehe nächste Seite

Fortsetzung Tabelle 1: Übersicht der bearbeiteten rekombinanten Proteine nichtribosomaler Peptidsynthetasen

Protein, Synthetase, Organismus	Vektor, Länge, Masse	Expression	Aufreinigung	Kristallisation	Optimierung	Diffraktion	pdb entry (Auflösung)
<b>Epimerisierungsdomänen</b>							
<b>FenD2 PCP-E</b> Fengycin Synthetase D <i>B. subtilis</i> F29-3	pQE-60 579 aa 66.8 kDa	✓	✓	✗	—	—	—
<b>TycA E</b> Tyrocidin Synthetase A <i>B. brevis</i> ATCC 8185	pQE-60 500 aa 56.8 kDa	✓	✓	✗	—	—	—
<b>TycA E (kurz)</b> Tyrocidin Synthetase A <i>B. brevis</i> ATCC 8185	pQE-60 466 aa 52.9 kDa	✓	✓	✓	✓	✓	nativ, 1.65 Å noch nicht veröffentlicht
<b>TycB3 E</b> Tyrocidin Synthetase B <i>B. brevis</i> ATCC 8185	pQE-60 487 aa 55.8 kDa	✓	✓	✓	✓	✓	nativer Datensatz bis 2.7 Å, Struktur bislang nicht gelöst
<b>TycB3 PCP-E</b> Tyrocidin Synthetase C <i>B. brevis</i> ATCC 8185	pQE-60 559 aa 63.9 kDa	✓	✓	✓	✗ <sup>★4</sup>	—	—
<b>GrsA PCP-E</b> Gramicidin Synthetase A <i>B. brevis</i> ATCC 9999	pQE-60 577 aa 67.6 kDa	✓	✓	✓	✗	—	—

**Anmerkungen:**

★4 Die erhaltenen kleinen, verwachsenen Kristalle konnten nicht optimiert werden.

**Referenzen zu den Ursprungsorganismen:**

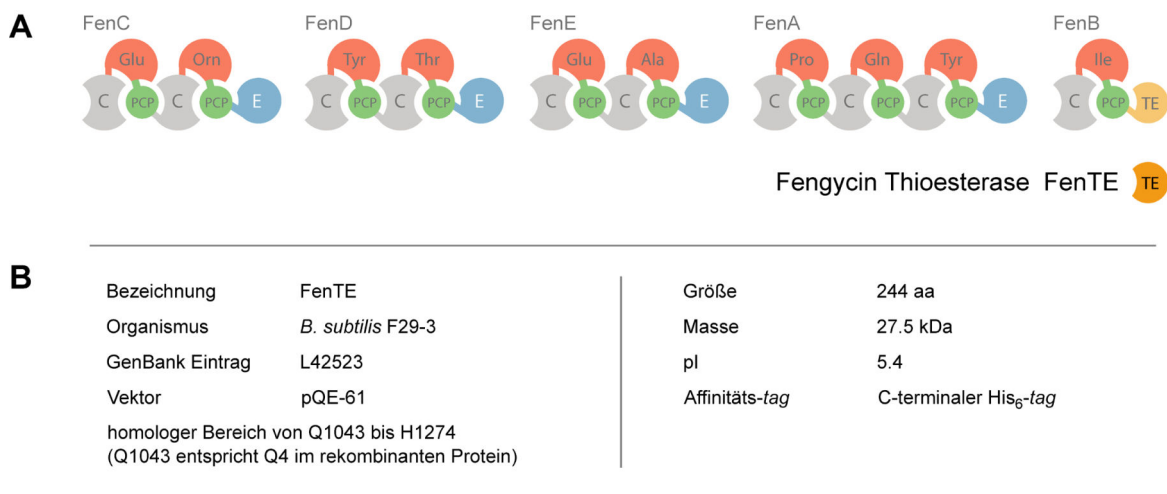
<i>Bacillus brevis</i> ATCC 8185	GenBank: AF004835	Mootz <i>et al.</i> , J. Bacteriol., <b>1997</b> , 179 (21), 6843-6850
<i>Bacillus brevis</i> ATCC 9999	GenBank: M29703	Kratzschmar <i>et al.</i> , J. Bacteriol., <b>1989</b> , 171 (10), 5422-5429
<i>Bacillus licheniformis</i> ATCC 10716	GenBank: AF007865	Konz <i>et al.</i> , Chem. Biol., <b>1997</b> , 4 (12), 927-937
<i>Bacillus subtilis</i> F29-3	GenBank: L42523	Chen <i>et al.</i> , Mol. Gen. Genet., <b>1995</b> , 248 (2), 121-125
<i>Stigmatella aurantiaca</i> DW4/3-1	GenBank: AAF19812	Silakowski <i>et al.</i> , J. Biol. Chem., <b>1999</b> , 274 (52), 37391-37399
<i>Streptomyces coelicolor</i> A3(2)	GenBank: NC_003888	Bentley <i>et al.</i> , Nature, <b>2002</b> , 417 (6885), 141-147

## 4.1 Struktur-Funktions-Beziehungen einer NRPS-Thioesterasedomäne

In den meisten nichtribosomalen Peptidsynthetaseclustern katalysieren Thioesterasedomänen (TE-Domänen) die Abspaltung des jeweiligen Produkts von der Synthetase. Diese Produktfreisetzung, die ggf. unter vorheriger Oligomerisierung der Peptidintermediate und Zyklisierung des Produkts erfolgt, stellt den letzten Schritt des Synthetaseclusters dar. Daher liegen die TE-Domänen, bezogen auf die Organisation des Clusters, an deren Ende. Im Falle des Fengycin Synthetaseclusters entspricht dies dem C-Terminus der Synthetase FenB (Abbildung 4.1A).

Allgemein bestehen Thioesterasedomänen aus etwa 240 Aminosäuren und gehören strukturell zu den  $\alpha/\beta$ -Hydrolasen. Das rekombinante Protein FenTE, die Typ I Thioesterase-Domäne des Fengycin Biosyntheseclusters aus *Bacillus subtilis* F29-3, besteht aus 244 Aminosäuren und wurde strukturell aufgeklärt [Samel, 2004] (Abbildung 4.1B).

In der vorliegenden Arbeit sollte der Bindungsmodus dieses makrozyklisierenden Enzyms untersucht werden. Dazu sollte zunächst die Struktur der mit Phenylmethylsulfonylfluorid (PMSF) kovalent inhibierten FenTE verfeinert werden. Des Weiteren sollten native Kristalle der FenTE mit substratanalogen Verbindungen umgesetzt werden (*Soaking*). Die so erhaltenen Komplexe von Enzym und enzymgebundener Verbindung sollten dann kristallographisch nachgewiesen werden und Aufschluss über Art der Wechselwirkungen sowie über die beteiligten Gruppen geben.



**Abbildung 4.1: Die Thioesterasedomäne des Fengycin Biosyntheseklusters**

(A) Die Thioesterasedomäne und ihre Position im Fengycin Biosynthesekluster aus *B. subtilis* F29-3.

(B) Angaben zur Synthetase FenB sowie zum rekombinanten Protein FenTE.

#### 4.1.1 Verfeinerung der Struktur der PMSF-inhibierten Fengycin Thioesterasedomäne

Im Rahmen der Diplomarbeit wurden native Kristalle der Fengycin Thioesterasedomäne (FenTE) mit Phenylmethylsulfonylfluorid (PMSF) behandelt und kristallographisch untersucht. Zur Lösung der Struktur wurde die Struktur der nativen FenTE verwendet. Die Verfeinerung dieser Struktur erfolgte jedoch erst zu Beginn der vorliegenden Arbeit. Die kristallographischen Details dieser Verfeinerung sind in Tabelle 2 wiedergegeben.

**Tabelle 2: Verfeinerungsstatistik zum Datensatz der PMSF-inhibierten Fengycin Thioesterasedomäne**

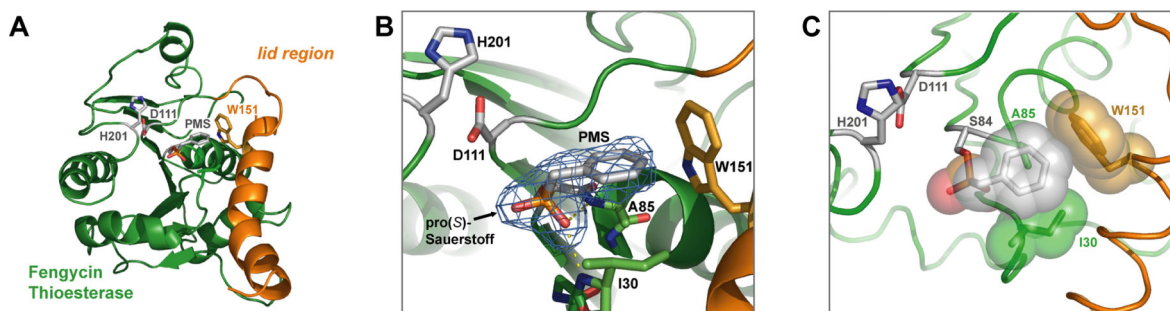
<b>Verfeinerung</b>	<b>PMS-FenTE Komplex</b>
Auflösung [Å]	24 - 2.5
Reflexionen (mit $F > 0\sigma$ )	8690
R-Faktor $^2$ ; freier R-Faktor (%)	16.2 (25.2); 23.7 (28.7)
Anzahl Proteinatome; Wassermoleküle; Heterogene	1665; 75; 10
r.m.s.d. der Bindungslängen [Å]	0.041
r.m.s.d. der Bindungswinkel (°)	2.589

**Erläuterungen:** Werte in Klammern beziehen sich stets auf die höchste Auflösungsschale.

$$2) R = \frac{\sum |F_{\text{obs}}| - k |F_{\text{calc}}|}{\sum |F_{\text{obs}}|} \times 100 \quad \text{mit dem Skalierungsfaktor } k.$$

In Abbildung 4.2A ist die verfeinerte Struktur der inhibierten FenTE dargestellt. Die kovalente Bindung der Phenylmethylsulfonyleinheit (PMS) wirkt sich lediglich auf die unmittelbare Umgebung aus und erzwingt eine strukturelle Veränderung der Seitenketten der Reste D111 und H201. Darüberhinaus ist das in der Struktur der nativen Fengycin Thioesterase beobachtete Acetation im Falle der inaktivierten Thioesterasedomäne aus dem aktiven Zentrum verdrängt. Die restliche Struktur der FenTE bleibt dagegen praktisch unverändert.

Abbildung 4.2B zeigt das aktive Zentrum der inaktivierten FenTE im Detail. Die Aminosäurereste I30 und A85 bilden das Oxyanionenloch. Im natürlichen Reaktionszyklus der Thioesterasedomäne stabilisieren sie die bei der Bindung und Abspaltung intermediär auftretende negative Ladung über Wasserstoffbrückenbindungen. Im Falle des Inhibitors treten sie mit dem pro(*R*)-Sauerstoff der Phenylmethylsulfonyleinheit in Wechselwirkung. Für den PMS-Anteil ist die Elektronendichte dargestellt, die zeigt, dass der Phenylrest in seiner Rotation eingeschränkt ist. Der Grund für diese bevorzugte Orientierung des Phenylrestes erschließt sich aus der Betrachtung seiner näheren Umgebung (Abbildung 4.2C). In der raumfüllenden Darstellung wird deutlich, dass der Phenylrest in einer hydrophoben Tasche gebunden vorliegt, die durch die Seitenketten der Reste I30, A85 und W151 gebildet wird. Die Seitenkette des Restes A85 liegt unterhalb des Phenylrestes und verhindert dessen freie Rotation. Aufgrund dieses Vorliegens einer hydrophoben Tasche in der Umgebung des Restes S84 wurde vermutet, dass der C-terminale Rest des natürlichen Substrats, L-Ile11, in dieser Tasche gebunden wird.



**Abbildung 4.2: Struktur der PMSF-inhibierten Fengycin Thioesterasedomäne**

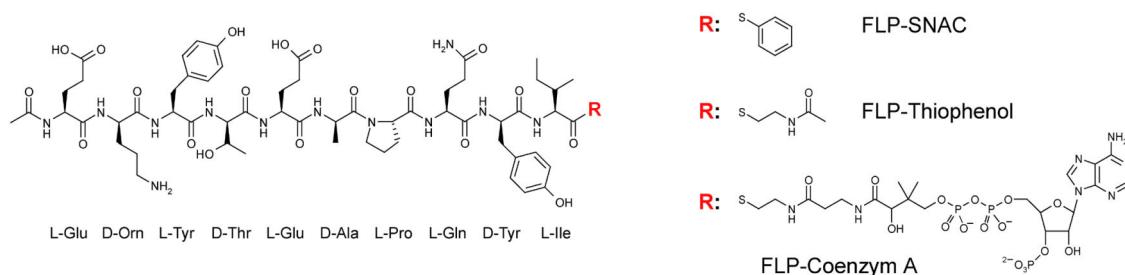
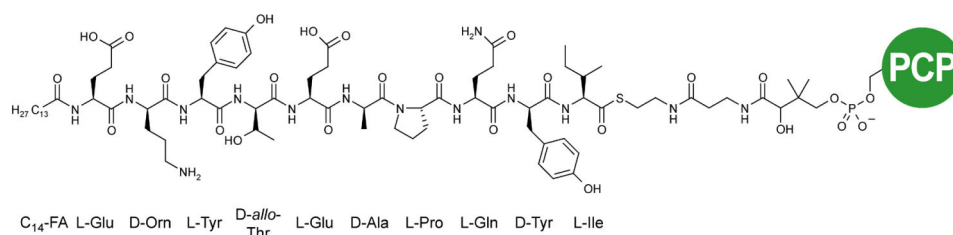
(A) Gesamtansicht der 2.5 Å Kristallstruktur. Der katalytisch aktive Serinrest S84 ist in dem inaktivierten Enzym kovalent modifiziert. (B) Detailansicht des aktiven Zentrums mit der Darstellung der Elektronendichte für die serin-gebundene Phenylmethysulfonyleinheit (PMS). Anhand der Elektronendichte zeigt sich eine Vorzugsorientierung für den Phenylrest des PMS-Anteils. I30 und A85 bilden das Oxyanionenloch. (C) Aufsicht auf das aktive Zentrum der FenTE mit ausgewählten Resten in raumfüllender Darstellung. Der PMS-Phenylrest ist in einer hydrophoben Tasche gebunden, wobei die Seitenkette des Restes A85 die Rotation der Phenylgruppe einschränkt.

#### 4.1.2 Soaking mit substratanalogen Verbindungen

Die substratanalogen Verbindungen wurden von der Arbeitsgruppe Prof. Marahiel zur Verfügung gestellt und waren dort für die biochemische Charakterisierung der Fengycin Thioesterasedomäne synthetisiert worden. Um das gegebenenfalls assoziierte Peptid kristallographisch im Enzym nachzuweisen, durfte es im Falle einer Einlagerung in die Bindungstasche nicht umgesetzt werden. Wie Sieber *et al.* gezeigt hatten, ist das an *S*-N-Acetylcysteamin (SNAC) gebundene Fengycin Lipopeptid (FLP) nicht ausreichend aktiviert und wird daher nicht von der Thioesterase umgesetzt [Sieber, 2003]. Daher eignete sich dieses Substratpeptid für *Soaks* über lange Zeiträume. Daneben wurden auch das reaktivere FLP-Thiophenol sowie das FLP-CoA für *Soaking*-Experimente eingesetzt. Eine Auswahl der Substratverbindungen zeigt Abbildung 4.3.

Des Weiteren wurde eine Variante des FLP-Thiophenol eingesetzt, bei dem das intramolekulare Nukleophil, Tyrosin 4, durch ein *para*-Aminophenylalanin ersetzt war. Untersuchungen mit der Fengycin TE-Domäne hatten ergeben, dass dieses Peptid – wie das reguläre FLP-Thiophenol – auf den Serinrest der Fengycin Thioesterase übertragen wird. Allerdings blieb die anschließende Zyklisierung aus, so dass keine Makrolaktamisierung, sondern ausschließlich die Hydrolyse der kovalenten Bindung des Enzym-Substrat-Addukts beobachtet wurde (Dr. Björn Wagner, Arbeitsgruppe Prof. Marahiel, persönliche Mitteilung). Da die beiden Thiophenol-Derivate von der TE-Domäne umgesetzt werden, wurden die Kristalle mit diesen Substraten nur kurz inkubiert. Damit sollte erreicht werden, dass sich das gebundene Substrat in der Bindungstasche nachweisen lässt, bevor es vollständig zum Produkt umgesetzt wurde.

## 4 Ergebnisse

C<sub>14</sub>-FLP-S-Ppan-PCP

**Abbildung 4.3: Auswahl der für *Soaking*-Experimente verwendeten, substratähnlichen Peptidverbindungen**

Neben den N-Acetylcysteamin-, Thiophenol- und CoA-Derivaten des synthetisierten, linearen Fengycin-Vorläuferpeptids ist zum Vergleich auch das natürliche, PCP-gebundene Substrat der Thioesterasedomäne dargestellt.

Kristalle des nativen Proteins wurden in Tropfen von Cryopuffer überführt, die die jeweilige Oligopeptid-Verbindung in Konzentrationen von 2 bis 10 mM enthielten. Sie wurden für die Dauer von einigen Minuten bis zu mehreren Tagen inkubiert und dann in flüssigem Stickstoff schockgefroren. Die Analyse der inkubierten Kristalle bei 2.2 bis 2.8 Å Auflösung ergab jedoch, dass in keinem der untersuchten Ansätze eine Verbindung im aktiven Zentrum der Thioesterase vorlag. Damit blieb allerdings offen, ob die Substrate nur nicht zum aktiven Zentrum gelangen konnten oder ob die Affinität zwischen Enzym und substratanaloger Verbindung zu gering war, um eine definierte Anlagerung zu bewirken.

### 4.1.3 **Modelling und Moleküldynamik-Simulation eines FenTE-Fengycin-Komplexes**

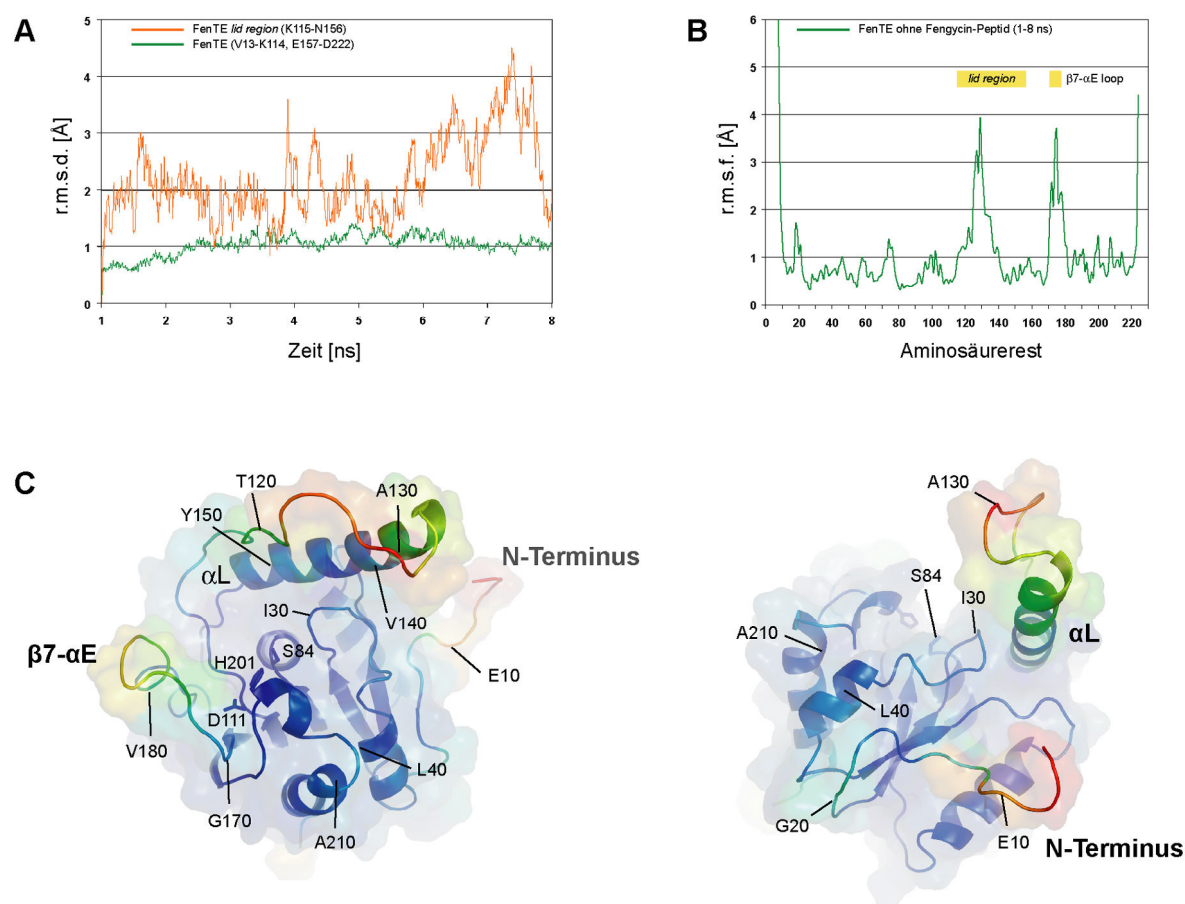
Nachdem die Versuche, einen Enzym-Substrat-Komplex herzustellen, fehlgeschlagen waren, sollte alternativ mit Hilfe computergestützter Verfahren ein Modell eines binären Komplexes aus der Fengycin TE-Domäne und dem Fengycin Lipopeptid berechnet werden. Als Vorbereitung darauf wurde zunächst die Fengycin Thioesterase in einer Moleküldynamik-Simulation auf ihre konformationelle Stabilität hin untersucht. Allerdings war in der Struktur der Fengycin Thioesterase ein Teil der *lid region* (T123-A129) in der Elektronendichtekarte nicht definiert. Um dennoch die Dynamik dieses Bereichs in den Berechnungen realistisch zu berücksichtigen, wurden die fehlenden Aminosäurereste mit dem Programm Swiss-Model [Schwede, 2003] ergänzt. Die erhaltene Proteinstruktur wurde anschließend in einer Moleküldynamik-Simulation eingesetzt. Nach Zugabe von  $\text{Na}^+$ -Ionen zum Ausgleich der überschüssigen negativen Ladungen im Protein wurde das Protein in einer mit Wasser gefüllten Zelle (*water box*) äquilibriert. Diese Simulation des unkomplexierten Proteins umfasste 4 Millionen Schritte von jeweils 2 fs, so dass die Gesamtdauer 8 ns betrug. Während dieses Zeitraums wurden im Bereich der *lid region*, abgesehen von der inhärenten Mobilität, keine gravierenden Änderungen beobachtet. Dies zeigte, dass die *lid region* ausreichend stabil ist und dass sie den Zugang zum aktiven Zentrum nicht beeinträchtigt. Damit war eine Voraussetzung für die anschließenden Arbeiten gegeben.

Unter Berücksichtigung der Tatsache, dass gerade zu Beginn gravierende Veränderungen auftreten können, wurde die Struktur des Komplexes nach einer Simulationsdauer von 1 ns als Grundlage für die weiteren Änderungen gewählt. Nach einer anfänglichen Äquilibration blieb der Kern der Fengycin Thioesterase (V13-K114 und E157-D222) über die gesamte Simulationsdauer hinweg weitgehend unverändert. Dies ist aus dem in Abbildung 4.4A grün dargestellten zeitlichen Verlauf der über diesen Proteinbereich gemittelten Abweichungen (r.m.s.d.) erkennbar. Die *lid region* (K115-N156) hingegen zeigte eine ausgeprägte Mobilität und wird daher hier separat betrachtet. Die Abweichungen sind in Abbildung 4.4A in rot wiedergegeben. Die erhöhte Flexibilität der *lid region* überrascht nicht, da die entsprechenden Reste in der Kristallstruktur neben dem  $\beta 7$ - $\alpha$ E-Loop die höchsten B-Faktoren aufweisen und zudem sieben Aminosäurereste (T123-A129) in der Elektronendichte gänzlich fehlen.

In Abbildung 4.4B sind die räumlichen Veränderungen (Fluktuationen, r.m.s.f.) der  $\text{C}_\alpha$ -Positionen des MD-Simulationslaufs dargestellt. Das Ausmaß der Fluktuationen wird für jedes  $\text{C}_\alpha$  durch die Farbe des Aminosäurerestes wiedergegeben (rot: r.m.s.f. = 4.7 Å, blau: r.m.s.f. = 0.4 Å). Hierbei fällt auch die hohe Beweglichkeit des  $\beta 7$ - $\alpha$ E-Loop auf. Diese Region gehört neben der *lid region* und dem N-Terminus zu den beweglichsten Bereichen des Proteins. Die hohe Mobilität des N-Terminus lässt sich damit erklären, dass dieser im rekombinanten Protein künstlich erzeugt wurde.



## 4 Ergebnisse



**Abbildung 4.4: Analyse der Moleküldynamik-Simulationen mit FenTE und dem Enzym-Produkt-Komplex**  
**(A)** Änderung der gemittelten Abweichungen (r.m.s.d.) gegenüber der Kristallstruktur in Abhängigkeit von der Zeit für die unkomplexierte Fengycin Thioesterase. Wegen der unterschiedlichen Mobilität sind *lid region* und Kernbereich separat dargestellt. **(B)** In der Simulation der unkomplexierten Fengycin Thioesterase beobachtete Fluktuationsverteilung (r.m.s.f.) der Hauptkettenatome (N, C $_{\alpha}$ , C). Neben den inhärent mobilen Termini lässt diese sequenzabhängige Auftragung zwei überdurchschnittlich bewegliche Regionen erkennen. Dabei handelt es sich um die *lid region* und den β7-αE-Loop. **(C)** Graphische Darstellung der Mobilität in der Simulation der unkomplexierten FenTE aus zwei verschiedenen Perspektiven.

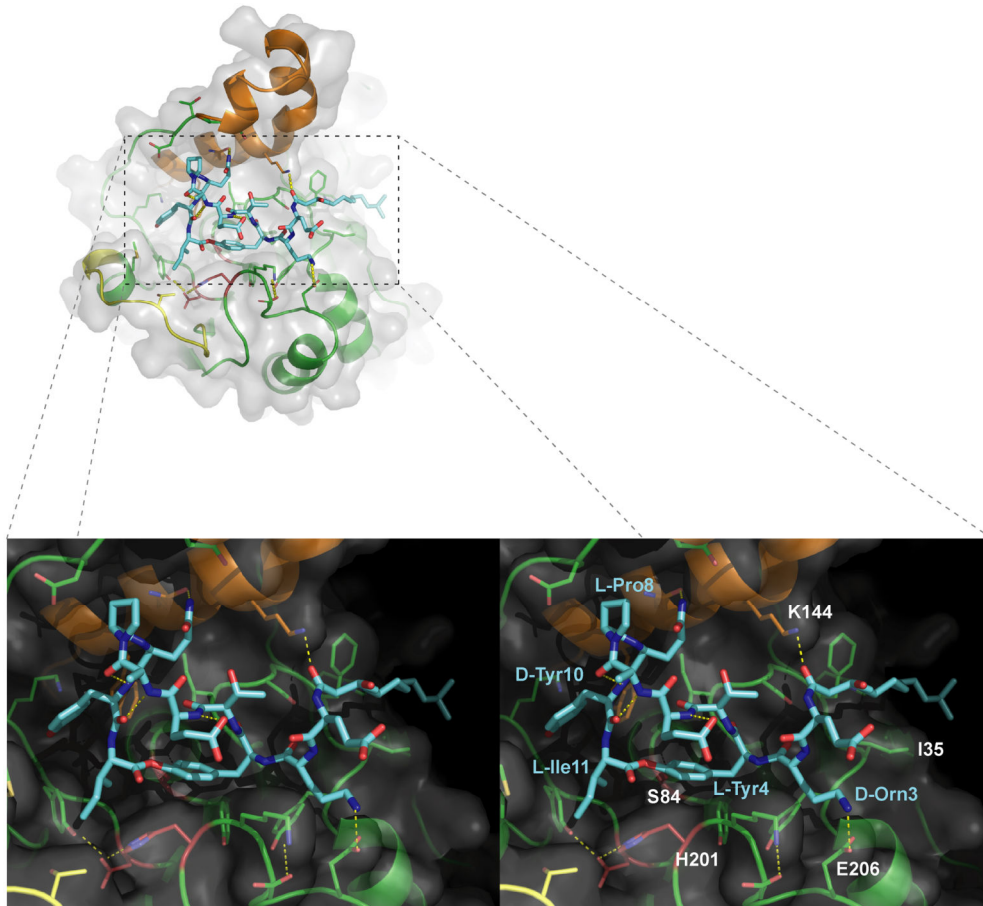
### 4.1.4 Docking und moleküldynamische Simulation des Enzym-Produkt-Komplexes

Für die nachfolgenden Experimente wurde von der Kristallstruktur der PMSF-inhibierten Fengycin Thioesterase (ohne den Inhibitoranteil) sowie von der NMR-Struktur des zyklischen Produktpeptids Fengycin [Volpon, 2000] ausgegangen.

Die Struktur des Fengycins wurde mit dem Programm SYBYL 6.7 (Tripos, Inc.) in die Bindungstasche des Enzyms FenTE eingepasst. Um einen biologisch relevanten Zustand darzustellen, musste der Komplex die Makrozyklisierung erlauben. Dazu sollte der Abstand zwischen den an der Zyklisierung beteiligten Resten, S84 und L-Ile11, nicht mehr als 3.5 Å betragen. Diese Bedingung grenzt die Anzahl der möglichen Orientierungen in der Bindungstasche deutlich ein. Prinzipiell verblieben zwei Orientierungen des Peptids, die die obige Bedingung erfüllen. Nur eine davon erschien jedoch biologisch plausibel und brachte die funktionellen



Gruppen des Produktpeptids in die Nähe komplementärer Gruppen des Enzyms. Anschließend wurde das Modell in einem Simulationslauf über eine halbe Nanosekunde äquilibriert. Die Gesamtdauer der Simulation des Komplexes betrug 10 ns. In Abbildung 4.5 ist die Struktur des FenTE-Fengycin-Komplexes nach 10 ns als Stereoabbildung dargestellt.



**Abbildung 4.5:** Stereoabbildung der Interaktionen des Lipopeptids Fengycin im aktiven Zentrum der Thioesterasedomäne FenTE

Eine detaillierte Betrachtung der Komplexstruktur sowie der Wechselwirkungen und der Aminosäurereste, die an der Substraterkennung beteiligt sind, findet im Rahmen der Diskussion in den Kapiteln 5.1 bis 5.2 statt.

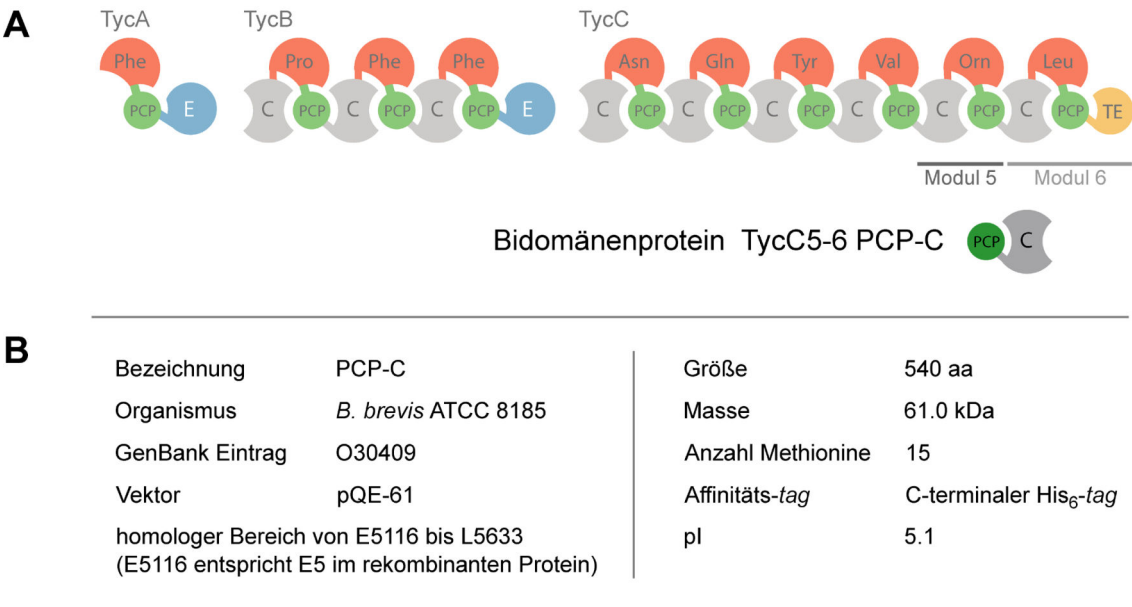
4.2 Kristallisation und Strukturlösung einer PCP-C Bidomäne

Das rekombinante Bidomänenprotein PCP-C wurde von Dr. Georg Schönafinger kloniert. Es stammt aus der Tyrocidin Synthetase C des Organismus *B. brevis* ATCC 8185. In der Tyrocidin Synthese erfüllen die beiden Domänen die nachfolgend beschriebenen Aufgaben.

Die zuvor aktivierte Aminosäure Ornithin (L-Orn9) wird an den Kofaktor der PCP-Domäne gebunden. Auf diese wird daraufhin das *upstream* synthetisierte Oktapeptidylintermediat übertragen. Das resultierende, PCP-gebundene Nonapeptidylintermediat wird in einer C-Domänen-katalysierten Reaktion auf die *downstream* gelegene, PCP-gebundene Aminosäure L-Leu10 übertragen. Die im rekombinanten Protein enthaltene PCP-Domäne ermöglicht sowohl die kovalente Bindung der Aminosäure Ornithin als auch die des Nonapeptidylintermediats. Die C-Domäne katalysiert die letzte Kondensationsreaktion in der Biosynthese des Tyrocidins.

Wie bereits in der Einleitung beschrieben, sind Elongationsmodule mit einer Domänenabfolge C-A-PCP definiert. Demnach besteht die PCP-C Bidomäne aus der PCP-Domäne des 5. Moduls und der Kondensationsdomäne des 6. Moduls der Synthetase TycC (Abbildung 4.6A). Der Sachverhalt, dass die beiden Domänen unterschiedlichen Modulen angehören, kommt auch in der Bezeichnung TycC5-6 PCP-C zum Ausdruck.

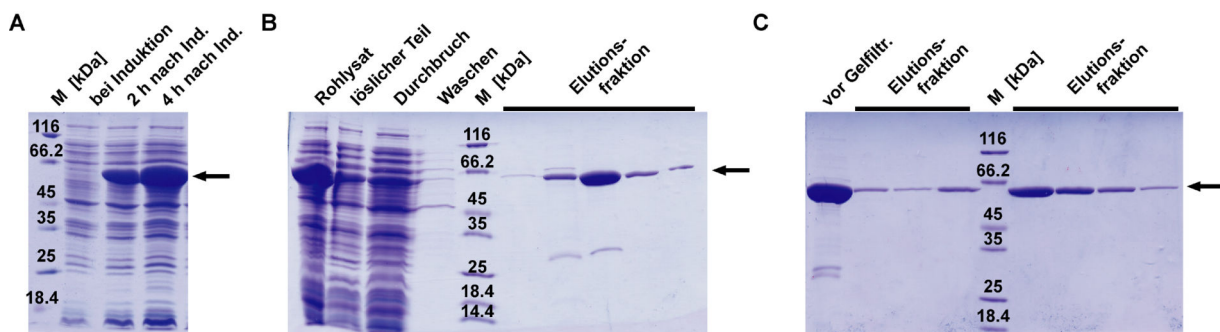
Bei der biochemischen Charakterisierung der PCP-C-Bidomäne konnte deren natürlich katalysierte Reaktion wegen der Instabilität des Substratpeptids nicht untersucht werden. Allerdings zeigte sie eine unerwartete Zyklisierungsaktivität gegenüber einem auf die PCP-Domäne übertragenen Hexapeptidylsubstrat [Schönafinger, 2003; Samel, 2007]. Das rekombinante Protein, welches in Abbildung 4.6B näher beschrieben ist, wurde in der vorliegenden Arbeit strukturell aufgeklärt.



**Abbildung 4.6: Das rekombinante Protein PCP-C aus dem Tyrocidin Synthetasecluster aus *B. brevis* ATCC8185**  
(A) Das rekombinante Protein PCP-C stammt aus der Tyrocidin Synthetase C. Die PCP-Domäne gehört zum 5. Modul, die Kondensationsdomäne zum 6. Modul der Synthetase TycC. (B) Informationen zum Ursprungsorganismus des Tyrocidin Synthetase Operons sowie einige Informationen zum rekombinanten Protein PCP-C.

#### 4.2.1 Expression und Aufreinigung der nativen PCP-C Bidomäne

Die Produktion des rekombinanten Proteins PCP-C erfolgte in LB-Medium bei 26 °C in einem Schüttelinkubator. Zur Induktion der Überproduktion wurde IPTG in einer Endkonzentration von 0.5 mM zugegeben. Abweichend von der zuvor beschriebenen Aufreinigung nativen PCP-Cs [Schönafinger, 2003] wurde hier im Anschluss an die Affinitätschromatographie eine Gelfiltration mit einer kalibrierten Superdex 200-Säule durchgeführt. Dies erlaubte anhand des ermittelten Elutionsvolumens die Abschätzung der Molekülmasse (siehe Kapitel 3.2.5). Für die PCP-C Bidomäne ergab sich eine apparente Masse von 67 kDa. Dieser Wert stimmt gut mit dem berechneten Molekulargewicht von 61 kDa überein und deutet darauf hin, dass PCP-C in Lösung monomer vorliegt. Die proteinhaltigen Fraktionen wurden vereinigt und auf eine Proteinkonzentration von 12 mg/mL aufkonzentriert. Die Ausbeute lag bei etwa 11 mg aufgereinigtes Protein PCP-C pro Liter LB-Expressionskultur.



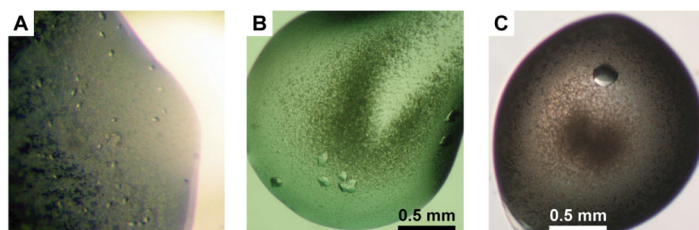
**Abbildung 4.7: 12% SDS-PAGE-Analyse der Produktion und Aufreinigung des Bidomänenproteins PCP-C**  
 Die Laufhöhe des rekombinanten Proteins ist jeweils mit einem Pfeil gekennzeichnet. (A) Zeitabhängige, induzierte Produktion der rekombinanten PCP-C-Bidomäne aus der Synthetase TycC. (B) Aufreinigung des Bidomänenproteins mittels Ni<sup>2+</sup>-NTA-Affinitätschromatographie. (C) Weitere Aufreinigung mittels Größenausschlusschromatographie.

#### 4.2.2 Kristallisation der nativen PCP-C Bidomäne

Bei einem anfänglich von Dr. Georg Schönafinger durchgeführten Kristallisationsscreen wurden Kristalle des rekombinanten Proteins PCP-C in einer Lösung aus 1.6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M MES (pH 6.5) und 10 % Dioxan beobachtet. Diese Kristalle konnten sowohl in *sitting drops* als auch in *hanging drops* reproduziert werden. Wie aus Abbildung 4.8A ersichtlich, entstanden unter der ursprünglichen Kristallisationsbedingung des kommerziellen Screens viele kleine Kristalle. Im Rahmen der Optimierung wurde auch *Micro-Seeding* angewendet, bei dem neue Kristallisationsansätze mit Fragmenten zertrümterter, kleiner Kristalle angeimpft werden. Diese Methode lieferte Kristalle wie sie in Abbildung 4.8B dargestellt sind. Schließlich gelang es auch ohne *Seeding* Einkristalle geeigneter Größe zu produzieren (Abbildung 4.8C). Die besten Kristalle wurden bei

## 4 Ergebnisse

19 °C aus *hanging drops* mit 1.6 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.04 M MES (pH 6.5) und 4 % Dioxan sowie mit einer Proteinkonzentration von 7.5 mg/mL erhalten. Vor der röntgenkristallographischen Untersuchung wurden die Kristalle kurz in einem zusätzlich 15 % Glycerin enthaltenden Kristallisationspuffer inkubiert und in flüssigem Stickstoff schockgefroren.



**Abbildung 4.8: Optimierung der Kristallisationsbedingung der nativen Bidomäne PCP-C**

(A) Kristalle aus einem Kristallisationsscreen. (B) Kristalle nach *Seeding*.

(C) Tropfen mit einzeltem Einkristall ohne *Seeding*.

### 4.2.3 Datensammlung von nativem PCP-C

Die tetragonalen Kristalle des nativen Proteins PCP-C wurden an der *beamline* BW7A am EMBL (DESY, Hamburg) bei 100 K untersucht. Es wurde ein Datensatz bis zu einer Auflösung von 1.8 Å aufgenommen, dessen Details in Tabelle 3 beschrieben sind. Aus den Zelldimensionen und aus dem Molekulargewicht des Proteins ließ sich schließen, dass in der asymmetrischen Einheitszelle ein einziges Polypeptid vorliegt (Matthews-Koeffizient  $V_M = 2.38 \text{ Å}^3/\text{Da}$ ; Solvensgehalt 42.5 %).

Nach Prozessierung mit dem Programm DENZO (HKL Research) wurde zunächst geprüft, ob sich die zur Strukturlösung erforderliche Phaseninformation mittels molekularen Ersatzes (*molecular replacement*) erhalten ließe. Als Grundlage für Suchmodelle diente hierbei die Struktur der isoliert stehenden Kondensationsdomäne VibH des Vibriobactin-Biosynthesecusters aus *Vibrio cholerae* [Keating, 2002]. Da die Sequenzidentität von VibH mit der Kondensationsdomäne aus PCP-C lediglich 19 % beträgt, war es nicht überraschend, dass mit Hilfe des *molecular replacement* keine Lösung gefunden werden konnte.

Zur Lösung des Phasenproblems sollte daher Selenomethionin-markiertes Protein PCP-C hergestellt und kristallisiert werden, das aufgrund der eingebrachten anomal streuenden Schweratome die experimentelle Bestimmung der Phasen durch MAD (*multiple anomalous dispersion*) ermöglichen sollte .

#### 4.2.4 Kristallisation von Selenomethionin-markiertem PCP-C

Anfängliche Probleme bei der Kristallisation des Selenomethionin-derivatisierten Proteins PCP-C ließen sich nach massenspektrometrischer Analyse darauf zurückführen, dass in dem zunächst produzierten rekombinanten Protein zwar alle Methionine durch Selenomethionin ersetzt waren; unbeabsichtigt lag aber auch die PCP-Domäne in der mit Phosphopantethein modifizierten *holo*-Form vor. Ursächlich hierfür war, dass das hierbei verwendete Minimalmedium keine Eisenionen enthielt. Aufgrund dieses Eisenmangels produzierten die Zellen des Expressionsstammes *E. coli* M15 [pREP4] das Siderophor Enterobactin über das zelleigene NRPS-Biosynthesecenter. Die mit diesen nichtribosomalen Peptidsynthetasen assoziierte Phosphopantethein-Transferase, EntD, modifizierte auch die PCP-Domäne des rekombinanten Proteins PCP-C.

Die unerwünschte posttranslationale Modifikation des rekombinanten Proteins konnte durch Zugabe von  $\text{FeSO}_4$  zum M9-Minimalmedium vermieden werden (siehe auch Kapitel 3.2.1). Die *apo*-Form des derivatisierten Proteins PCP-C kristallisierte unter vergleichbaren Bedingungen wie das native Protein, wobei der beste untersuchte Kristall von Selenomethionin-markiertem PCP-C in einem Tropfen aus 1.6 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.15 M MES (pH 6.2) und 4 % Dioxan entstand.

#### 4.2.5 MAD-Datensammlung von Selenomethionin-markiertem PCP-C

Die Datensammlung an einem Kristall von Selenomethionin-markiertem PCP-C erfolgte an der *beamline* BW7A am EMBL (DESY, Hamburg). Die drei verwendeten Wellenlängen wurden zuvor durch Aufnahme eines Fluoreszenzspektrums des Kristalls ermittelt. Die Datensätze sind in Tabelle 3 auf Seite 62 zusammengefasst.

#### 4.2.6 Strukturlösung der Bldomäne PCP-C

Zur Lösung der Struktur wurde das automatisierte Strukturlösungsprogramm AUTO-RICKSHAW [Panjikar, 2005] am EMBL-Hamburg genutzt. Unter Verwendung der MAD-Daten aus den drei Datensätzen des Kristalls von Selenomethionin-markiertem PCP-C wurden von dem Programm SHELX die Positionen von 13 der insgesamt 15 Selenatome gefunden. Die daraus erhaltene Elektronendichtekarte konnte durch Modifikation der Elektronendichte (*solvent flattening*) mit dem Programm DM [CCP4, 1994] weiter verbessert werden. Schließlich wurde mit ARP6.1 [Perrakis, 2001] ein erstes Modell in die Elektronendichte eingepasst.

## 4 Ergebnisse

**Tabelle 3: Datensammlung nativer und SeMet-substituierter Kristalle und Verfeinerung der Struktur von PCP-C**

Datensammlung	PCP-C (nativ)	PCP-C (Selenomethionin)		
		<i>peak</i>	<i>inflection</i>	<i>remote</i>
Wellenlänge $\lambda$ [Å]	0.808	0.9777	0.9780	0.9774
Raumgruppe	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2	P4 <sub>3</sub> 2 <sub>1</sub> 2
Zellparameter [Å]	a=b=84.97 c=164.97	a=b=85.09 c=164.76	a=b=85.05 c=164.69	a=b=85.17 c=164.91
Auflösung [Å]	20-1.85	25-2.50	25-2.50	25-2.50
beobachtete Reflexionen	261739	90992	90919	90144
einzigartige Reflexionen	51834	21485	21484	21555
I/ $\sigma$ (I)	20.9 (2.2)	18.1 (5.1)	18.0 (4.9)	17.9 (4.8)
R <sub>merge</sub> <sup>1</sup>	0.043 (0.582)	0.066 (0.234)	0.050 (0.247)	0.060 (0.250)
Vollständigkeit (%)	98.9 (100.0)	97.9 (99.0)	98.0 (99.2)	97.9 (99.0)

Verfeinerung	PCP-C (nativ)
Auflösung [Å]	19.8 - 1.85
Reflexionen (mit F>0 $\sigma$ )	49809
R-Faktor <sup>2</sup> ; freier R-Faktor (%)	20.8 (26.4); 23.8 (29.7)
Anzahl Proteinatome; Wassermoleküle; Ionen; Heterogene	4160; 294; 1; 22
r.m.s.d. der Bindungslängen [Å]	0.006
r.m.s.d. der Bindungswinkel (°)	1.094

**Erläuterungen:** Werte in Klammern beziehen sich stets auf die höchste Auflösungsschale.

$$1) R_{\text{merge}} = \frac{\sum \sum |I_j(h) - \langle I(h) \rangle|}{\sum \sum I_j(h)} \quad 2) R = \frac{\sum | |F_{\text{obs}}| - k |F_{\text{calc}}| |}{\sum |F_{\text{obs}}|} \times 100 \quad \text{mit dem Skalierungsfaktor } k.$$



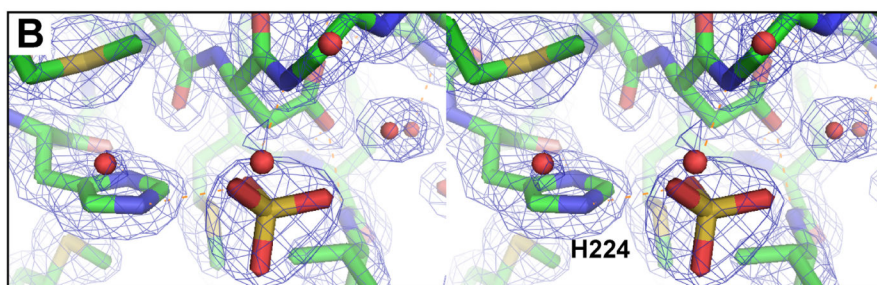
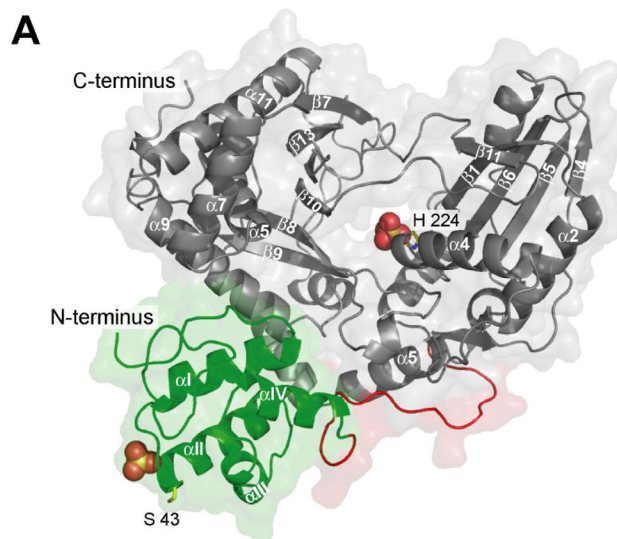
Anschließend wurde das Strukturmodell manuell mit COOT [Emsley, 2004] vervollständigt und mit REFMAC5 [CCP4, 1994] bis zu einem R-Faktor von 20.8 % (freier R-Faktor 23.8 %) verfeinert. Die kristallographischen Details der Verfeinerung sind in Tabelle 3 zusammengetragen.

In der Kristallstruktur ist die PCP-C Bidomäne von R3 bis M455 und von Q461 bis L522 definiert. Daneben enthält die asymmetrische Einheitszelle 294 Wassermoleküle sowie zwei Sulfationen und zwei Dioxanmoleküle, die aus dem Kristallisationspuffer stammen. Die Kristallstruktur von PCP-C ist in Abbildung 4.9 dargestellt und zeigt die PCP-Domäne (M1-T82) in grün und die größere Kondensationsdomäne (V101-L522) in grau. Beide Domänen sind durch den rot dargestellten Linker (A83-P100) verbunden. Eine eingehende Beschreibung der Struktur folgt im Rahmen der Diskussion (Kapitel 5.3 bis 5.5).

**Abbildung 4.9: 1.8 Å Kristallstruktur der PCP-C-Bidomäne der Tyrocidin Synthetase C aus *B. brevis* ATCC 8185**

**(A)** Gesamtansicht des Bidomänenproteins PCP-C mit den beiden nahe den jeweiligen aktiven Zentren gebundenen Sulfationen.

**(B)** Stereoabbildung des aktiven Zentrums der Kondensationsdomäne mit Darstellung der  $2F_0 - F_c$  Elektronendichte (Konturierung 1.5  $\sigma$ ). Wie in der Diskussion ausführlich beschrieben, integriert das Sulfation mit mehreren Aminosäureresten der Kondensationsdomäne.

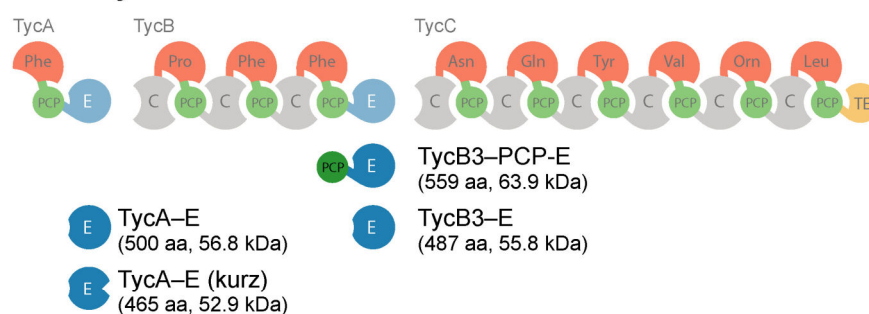


### 4.3 Arbeiten zur Strukturaufklärung von Epimerisierungsdomänen

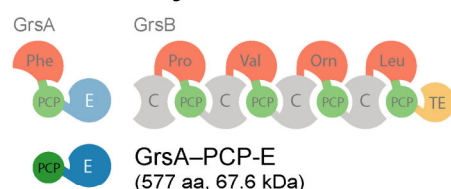
Im Rahmen dieser Arbeit wurden mehrere Epimerisierungsdomänen auf ihre Eignung zur Kristallisation hin untersucht. Basierend auf den Arbeiten in der Arbeitsgruppe Prof. Marahiel lagen bereits einige rekombinante Proteine als Plasmid vor. Diese wurden in *E. coli* überexprimiert, aufgereinigt und mit Hilfe der Kristallisationsscreens getestet.

Bei den rekombinanten Proteinen handelte es sich entweder um einzelne Epimerisierungsdomänen (E) oder um PCP-E-Bidomänen (PCP-E). Die kodierenden Sequenzen waren zuvor in der Arbeitsgruppe Prof. Marahiel mittels PCR aus der genomischen DNA der jeweiligen *Bacillus*-Produzentenstämme amplifiziert und in pQE-Vektoren kloniert worden. Mit Ausnahme der C-terminal verkürzten Epimerisierungsdomäne aus TycA (TycA-E kurz) lagen zu Beginn der

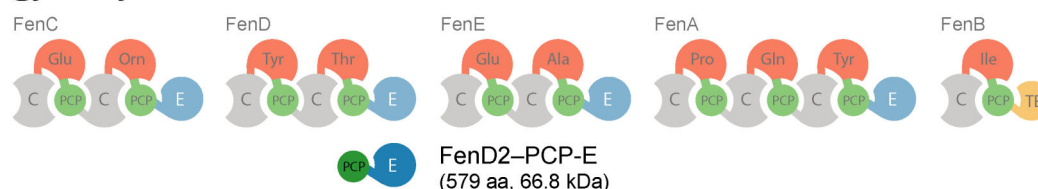
#### Tyrocidin Synthetase Cluster aus *Bacillus brevis* ATCC 8185



#### Gramicidin S Synthetase Cluster aus *Bacillus brevis* ATCC 9999



#### Fengycin Synthetase Cluster aus *Bacillus subtilis* F29-3



#### Abbildung 4.10: Übersicht der bearbeiteten Epimerisierungsdomänen

Die Graphik veranschaulicht den Bezug zwischen den rekombinanten Proteinen E bzw. PCP-E und ihren zugehörigen Biosyntheseclustern. Für jedes untersuchte Protein sind die Anzahl der Aminosäuren und die Masse angegeben.



Arbeit alle Plasmide vor. Abbildung 4.10 zeigt alle rekombinanten Proteine und gibt Aufschluss über die Position der Epimerisierungsdomänen im Kontext der zugehörigen Biosynthesecuster. Im folgenden Teil dieses Kapitels werden zunächst unter 4.3.1 die Ergebnisse und die Kristallstruktur der verkürzten Epimerisierungsdomäne aus der Tyrocidin Synthetase A vorgestellt. Unter 4.3.2 folgen dann die Ergebnisse der kristallographischen Arbeiten mit der Epimerisierungsdomäne aus der Tyrocidin Synthetase B. Zwar wurden von diesem Protein Kristalle erhalten, die die Aufnahme prozessierbarer Datensätze erlaubten, jedoch konnten bislang die zur Strukturlösung erforderlichen Phasen nicht bestimmt und die Struktur daher nicht gelöst werden.

Die weiteren in Abbildung 4.10 dargestellten rekombinanten Proteine (TycA-E, TycB3-PCP-E, GrsA-PCP-E und FenD2-PCP-E) lieferten keine oder keine kristallographisch brauchbaren Kristalle. Die Ergebnisse der Arbeiten mit diesen vier Proteinen werden daher zusammen unter 4.3.3 behandelt.

#### **4.3.1 Die C-terminal verkürzte Epimerisierungsdomäne der Tyrocidin Synthetase A**

##### **Klonierung und Expression von *tycA-E* (kurz)**

Die Verkürzungsmutante, der die letzten 35 Aminosäurereste der E-Domäne fehlen, die aber den vektor-kodierten His<sub>6</sub>-tag enthält, wurde mittels inverser PCR aus dem Plasmid pQE60\_*tycA-E* hergestellt. Die Methode der inversen PCR ist in Kapitel 3.1.3 beschrieben. Vor der Synthese der DNA-Primer wurde über den Xtal-Pred-Server [Slabinski, 2007] (<http://ffas.burnham.org/XtalPred-cgi/xtal.pl>, Februar 2008) eine Sekundärstrukturvorhersage für das ursprünglich kodierte, rekombinante Protein TycA-E mit dem Programm PSIPRED [Jones, 1999] durchgeführt. Der C-Terminus des rekombinanten Proteins mit seinen vorhergesagten Sekundärstrukturelementen ist in Abbildung 4.11 dargestellt. Zum Vergleich ist in Abbildung 4.11A die natürliche Aminosäuresequenz aus *Bacillus brevis* ATCC 8185 angegeben, die die Grundlage für die frühere Klonierung bildete [Mootz, 1997]. Aufgrund der engen Verwandtschaft zwischen Epimerisierungs- und Kondensationsdomänen wurde das Ergebnis der Sekundärstrukturvorhersage mit den bekannten Strukturen der Kondensationsdomänen VibH und PCP-C verglichen. Die vorhergesagten, sekundären Strukturelemente ließen sich denen in VibH und PCP-C gut zuordnen. Die entsprechenden C-Termini von VibH und PCP-C sind mit ihrer experimentell bestimmten Sekundärstruktur in Abbildung 4.11D und E wiedergegeben. Die vorhergesagte, 20 Aminosäurereste umfassende  $\alpha$ -Helix in TycA-E entsprach der  $\alpha$ -Helix  $\alpha$ 11 in VibH und PCP-C. Die PCR-Primer wurden so gewählt, dass in der verkürzten TycA-Epimerisierungsdomäne die Helix komplett erhalten bleibt und der His<sub>6</sub>-tag sich daran anschließt (Abbildung 4.11C).

## 4 Ergebnisse

<b>A Sequenz TycA</b> Genbank AF004835 (Mootz und Marahiel, 1997)	1027	SEQY REESIQQLSQ SYQKHLIAII AHCTEKKEVE RTPSDFSVKG LQMEEMDDIF ELLANTLR	1088
<b>B vorhergesagte Sekundärstruktur TycA-E</b> rekombinantes Protein (Linne et al., 2001)	430	SEQY <b>REESIQQLSQ SYQKHLIAII AHCTEKKEVE</b> <b>RT</b> PSDFSVKG <b>LQMEEMDDIF ELLANTLRGS</b> RSHHHHHH	500
<b>C Sequenz der verkürzten TycA-E</b> rekombinantes Protein in dieser Arbeit hergestellt	430	SEQY REESIQQLSQ SYQKHLIAII AHC-----L QSHHHHHH	464
<b>D beobachtete Sekundärstruktur TycC6-C</b> PDB Eintrag: 2JGP (Samel et al., 2007)	486	TSLF <b>TRETMERFSR HFLTIAASIV QNP</b> HIRLGEI DML	522
<b>E beobachtete Sekundärstruktur VibH</b> PDB Eintrag: 1L5A (Keating et al., 2002)	382	LASY <b>PQSHWQSHCE RFP</b> RFFEQLL <b>ARFQ</b> QVEQDV <b>ARLLA</b> EPAA	424

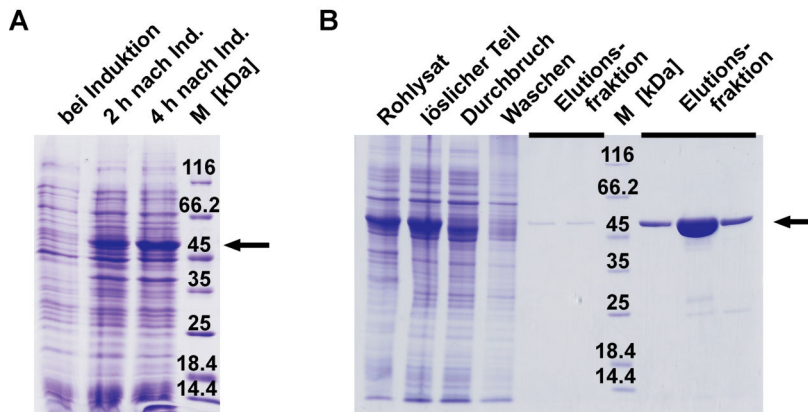
**Abbildung 4.11: Entfernung des C-terminalen COM-Bereichs unter Berücksichtigung der vorhergesagten Sekundärstruktur in TycA-E und der beobachteten Sekundärstruktur in NRPS-Kondensationsdomänen**

(A) Ausgangssequenz für alle besprochenen Arbeiten mit der TycA-E-Domäne ist das Tyrocidin-Operon aus *Bacillus brevis* ATCC 8185. (B) C-Terminus mit Sekundärstrukturvorhersage des ursprünglichen, rekombinanten Proteins TycA-E. Rote Buchstaben kennzeichnen  $\alpha$ -helikale Bereiche. (C) Sequenz des in dieser Arbeit klonierten, verkürzten TycA-E. (D) und (E) Vergleichbare Sequenz der C-Termini von PCP-C und VibH mit Angabe der dort beobachteten Sekundärstruktur.

Die Überlegungen, die ursprünglich zur Verkürzung der Epimerisierungsdomäne führten, sind im Anhang (Kapitel 8.5, Seite 135) dargelegt.

Das Produkt der inversen PCR wurde mittels Agarosegel überprüft. Es zeigte nur eine Bande auf der erwarteten Laufhöhe. Nach selektivem Verdau der Templat-DNA mit *DpnI* wurde das Amplifikat gereinigt. Die mit beiden DNA-Primern neu eingeführten *PstI*-Restriktionsschnittstellen wurden mit *PstI* verdaut, anschließend wurde das Plasmid religiert. Nach Transformation in den Klonierungsstamm DH5 $\alpha$  wurde das Plasmid präpariert und in den Expressionstamm *E. coli* M15 [pREP4] transformiert. Die DNA-Sequenz wurde per Didesoxy-Sequenzierung kontrolliert und die Molekülmasse des kodierten, rekombinanten Proteins durch die massenspektrometrische Analyse bestätigt (berechnete Masse: 52 782.4 Da; gemessene Masse:  $[M+H]^+=52\ 784$  Da). Dieser Proteinmasse nach ist im aufgereinigten Protein das N-terminale Formylmethionin abgespalten.

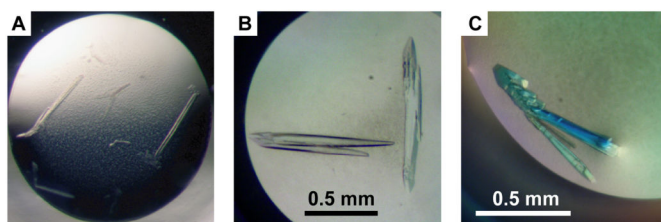
Die verkürzte TycA-Epimerisierungsdomäne wurde in LB Medium bei 30 °C durch Induktion mit 0.5 mM IPTG überproduziert. Anhand von drei Expressionen im analytischen Maßstab, mit denen der Einfluss der Induktorkonzentration auf die Löslichkeit des rekombinanten Proteins abgeschätzt werden konnte, zeigte sich, dass unter dieser Bedingung ausreichende Mengen des Proteins mit guter Löslichkeit erhalten wurden. Die Zellen wurden 4 Stunden nach Induktion abzentrifugiert und in Puffer A resuspendiert. Die Reinigung mittels Ni<sup>2+</sup>-NTA-Affinitätschromatographie lieferte die Epimerisierungsdomäne bereits in guter Reinheit. Die proteinhaltigen Fraktionen wurden vereinigt und mit Hilfe von Konzentratoren (Ausschlussgröße 30 kDa) aufkonzentriert. Durch Reinigung über eine PD-10-Säule wurde das restliche Imidazol entfernt, so dass das Protein im Gelfiltrationspuffer vorlag. Die das Protein enthaltenden Fraktionen wurden vereinigt und aufkonzentriert.



**Abbildung 4.12: 12% SDS-PAGE-Analyse der Überproduktion und Aufreinigung von TycA-E (kurz)**  
 (A) Zeitliche Abhängigkeit der induzierten Überproduktion der verkürzten TycA-Epimerisierungsdomäne.  
 (B) Aufreinigung von TycA-E per  $\text{Ni}^{2+}$ -NTA-Affinitätschromatographie.

### Kristallisation der verkürzten Epimerisierungsdomäne der Tyrocidin Synthetase A

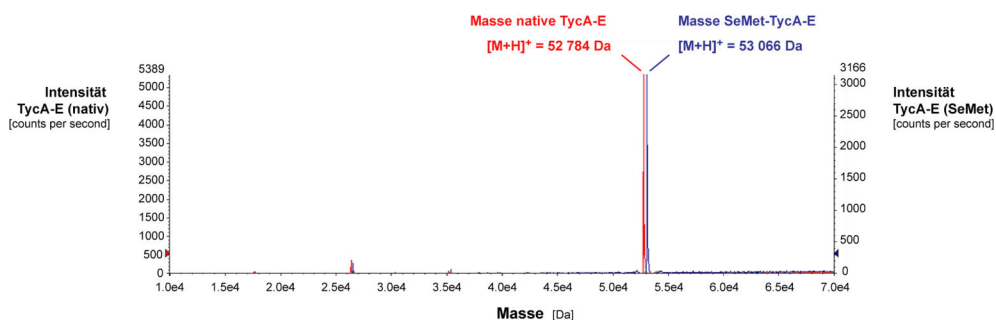
Für erste Kristallisationsversuche mit kommerziellen Screens der Firmen Sigma und Qiagen wurden zwei steril filtrierte Lösungen der verkürzten Epimerisierungsdomäne verwendet. Dabei wurde das Protein in zwei Konzentrationen, in 10 mg/mL und in 6 mg/mL, eingesetzt. An einem Cartesian Kristallisationsroboter wurden im 96-well-Format für jede Screenbedingung und für beide Proteinkonzentrationen jeweils ein *sitting drop*-Ansatz (Tropfengröße 300 nL + 300 nL) präpariert. Die Kristallisationsplatte wurde anschließend mit einer transparenten Folie luftdicht verschlossen und bei 19 °C unter Lichtausschluss inkubiert. Nach 4 Tagen zeigten sich unter verschiedenen Kristallisationsbedingungen farblose, nadel- bzw. säulenförmige Kristalle (siehe Abbildung 4.13A). Die Ansätze dieser Bedingung wurden mit selbstangesetzten Lösungen wiederholt. Die Kristalle, die aus 0.1 M Kaliumphosphat und 20 % PEG 8000 entstanden, waren am besten zu reproduzieren und schienen für eine Optimierung am besten geeignet. Die erste Optimierung umfasste eine Variation der Kaliumphosphatkonzentration, des pH-Werts des Kaliumphosphatpuffers, der PEG 8000-Konzentration sowie des PEG-Molekulargewichts. In einer zweiten Optimierung wurde auch der Einfluss der Temperatur auf die Kristallisation untersucht.



**Abbildung 4.13: Optimierung der Kristalle der verkürzten TycA-Epimerisierungsdomäne**  
 (A) Mehrere kleine Kristalle aus einem Screen im 96-well-Format (Tropfengröße 600 nL).  
 (B) und (C) Größere Kristalle aus einer manuellen Optimierung (Tropfengröße 2 µL). Die farblosen Kristalle sind dichroit und erscheinen bei Verwendung gekreuzter Polfilter farbig.

### Produktion der Selenomethionin-markierten Epimerisierungsdomäne aus TycA

Zur Strukturlösung war wiederum die experimentelle Bestimmung der Phasen erforderlich. Für eine Anwendung der MAD-Methode wurde Selenomethionin-markiertes Protein der verkürzten Epimerisierungsdomäne aus TycA hergestellt und analog zum nativen Protein aufgereinigt. Die quantitative Substitution der sechs Methionine durch Selenomethionin wurde durch Massenspektrometrie bestätigt (gemessene Proteinmasse:  $[M+H]^+ = 53\,066$  Da).



**Abbildung 4.14:** MS-Analyse der nativen und der Selenomethionin-markierten Proben der verkürzten TycA-E. Die Massendifferenz von 282 Da entspricht einer Substitution aller sechs Methione.

### Kristallisation der Selenomethionin-markierten Epimerisierungsdomäne aus TycA

Da sich die Kristallisationsbedingung des Selenomethionin-markierten Proteins von der des nativen Proteins unterscheiden kann, wurden bei 4 °C und 19 °C jeweils ein Screen, basierend auf der Kristallisationsbedingung des nativen Proteins, angefertigt. Dabei wurde in einem Screen sowohl die Konzentration und der pH-Wert der Pufferkomponente als auch die Konzentration und die Art des Präzipitans PEG variiert. Der beste Kristall wurde aus einem Ansatz mit einem Mischungsverhältnis von 1:2 (Proteinlösung: Kristallisationspuffer) und einer Lösung aus 0.05 M Kaliumphosphat und 22 % PEG 8000 erhalten.

### Datensammmlung der Epimerisierungsdomäne aus Tyrocidin Synthetase A

Zum Einfrieren dieser Kristalle wurden sie in einen Puffer aus 0.05 M Kaliumphosphat, 22 % PEG 8000 und 20 % Glycerin überführt und anschließend in flüssigem Stickstoff schockgefroren. Die am besten geordneten Kristalle zeigten am hausinternen Diffraktometer Röntgenbeugung mit einer maximalen Auflösung von 2.7 Å und Mosaizitäten um 0.6 °. Durch die Aufnahme erster Datensätze und die Indizierung wurde der Kristallform die Raumgruppe  $P2_12_12_1$  zugewiesen. Anhand der Zellparameter (siehe Tabelle 4) und einer Molekülmasse von 52.7 kDa konnte man davon ausgehen, dass die asymmetrische Einheitszelle ein Proteinmolekül enthält (Matthews-Koeffizient  $V_M = 2.08 \text{ Å}^3/\text{Da}$ ; Solvensgehalt 40.9 %). Von diesen Kristallen wurden an

der beamline ID 29 (ESRF, Grenoble) Datensätze mit einer maximalen Auflösung von 1.65 Å aufgenommen.

Die vorläufige Charakterisierung von Kristallen der Selenomethionin-substituierten TycA-E-Domäne am hausinternen Diffraktometer ergab, dass die Kristalle bis unter 3 Å diffraktierten. Die Aufnahme von MAD Datensätzen bei drei Wellenlängen wurde von Dr. Andreas Heine (Arbeitsgruppe Prof. Klebe, Fachbereich Pharmazie, Philipps-Universität Marburg) an der beam-

**Tabelle 4: Datensammlung nativer und SeMet-substituierter Kristalle und Verfeinerungsstatistik zu TycA-E**

Datensammlung	TycA-E (nativ)	TycA-E (Selenomethionin)		
		<i>peak</i>	<i>inflection</i>	<i>remote</i>
Wellenlänge $\lambda$ [Å]	1.039	0.97965	0.97984	0.91841
Raumgruppe	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Zellparameter [Å]	a=47.71 b=74.59 c=124.75	a=48.40 b=74.40 c=123.96	a=48.41 b=74.41 c=124.00	a=48.48 b=74.59 c=124.23
Auflösung [Å]	47.8-1.65	36.1-1.87	36.1-1.87	34.0-1.75
beobachtete Reflexionen	226421	464865	473060	552875
einzigartige Reflexionen	54352	38150	38184	46465
$I/\sigma(I)$	9.6 (1.7)	7.5 (2.0)	7.9 (2.2)	6.8 (3.1)
$R_{\text{merge}}^1$	0.058 (0.440)	0.055 (0.173)	0.055 (0.190)	0.059 (0.242)
Vollständigkeit (%)	99.9 (100)	100 (100)	100 (100)	100 (100)

Verfeinerung	TycA-E (nativ)
Auflösung [Å]	64.0 - 1.65
Reflexionen (mit $F > 0\sigma$ )	54053
R-Faktor <sup>2</sup> ; freier R-Faktor (%)	18.0 (24.7); 21.2 (25.8)
Anzahl Proteinatome; Wassermoleküle	4160; 399
r.m.s.d. der Bindungslängen [Å]	0.012
r.m.s.d. der Bindungswinkel (°)	1.337

**Erläuterungen:** Werte in Klammern beziehen sich stets auf die höchste Auflösungsschale.

$$1) R_{\text{merge}} = \frac{\sum \sum |I_j(h) - \langle I(h) \rangle|}{\sum \sum I_j(h)} \quad 2) R = \frac{\sum | |F_{\text{obs}}| - k |F_{\text{calc}}| |}{\sum |F_{\text{obs}}|} \times 100 \quad \text{mit dem Skalierungsfaktor } k.$$

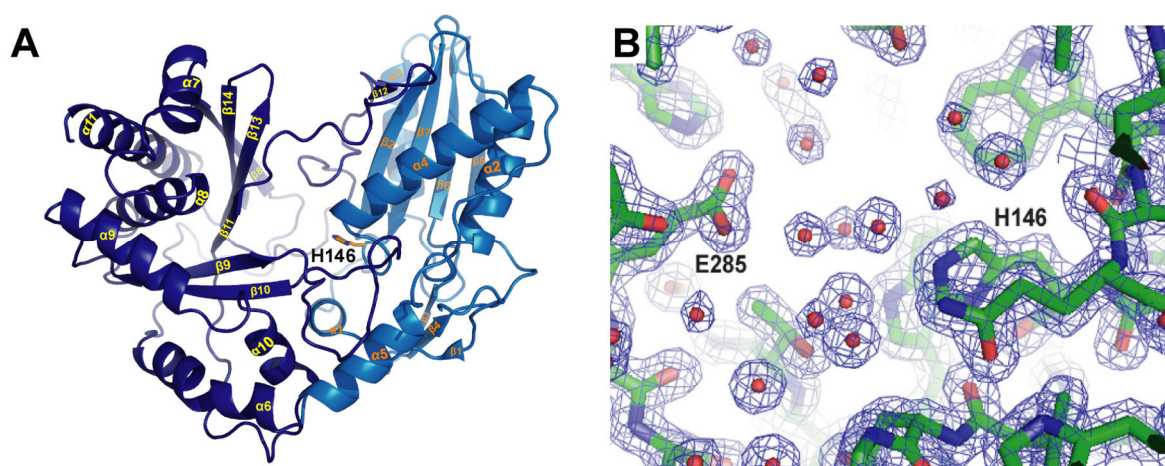
line 14.2 der Proteinstrukturfabrik (BESSY, Berlin) durchgeführt. Die intensive Synchrotronstrahlung ermöglichte die Datensammlung bis zu einer Auflösung von 1.85 Å.

### Strukturlösung der TycA-Epimerisierungsdomäne

Die Strukturlösung erfolgte erneut mit Hilfe der automatisierten Strukturlösungsplattform AUTORICKSHAW am EMBL, Hamburg [Panjikar, 2005]. Für die experimentelle Phasenbestimmung wurden drei Datensätze des Selenomethionin-substituierten Proteins TycA-E herangezogen.

Die Positionen der 6 Selenatome wurden mit dem Programm SHELXD ermittelt [Schneider, 2002]. Nach Modifikation der Elektronendichte wurden die Phasen mit dem Programm SHELXE [Sheldrick, 2002] berechnet. Schließlich wurde mit dem Programm ARP/wARP [Perrakis, 1999] ein Modell in die Elektronendichte eingepasst, das bereits 80 % der Struktur beschrieb. Die Struktur des nativen Proteins wurde mit dem Programm REFMAC 5 [CCP4, 1994] weiter verfeinert. Am Ende der Verfeinerung lagen der R-Faktor bei 18.0 % und der freie R-Faktor bei 21.2 %.

Die Epimerisierungsdomäne der Tyrocidin Synthetase A stellt die erste NRPS-Epimerisierungsdomäne dar, deren Struktur gelöst wurde. Die 1.65 Å Kristallstruktur des 464 Aminosäuren umfassenden, rekombinanten Proteins ist von den Resten K7 bis H460 definiert und enthält 399 Wassermoleküle. Sie besteht wie die NRPS-Kondensationsdomänen aus zwei Subdomänen, die jeweils das Faltungsmotiv der Chloramphenicol-Acetyltransferase zeigen. Die Struktur der Epimerisierungsdomäne und die Implikationen für die Epimerisierung PCP-gebundener Aminosäuren werden in der Diskussion eingehend behandelt (Kapitel 5.6 bis 5.7).



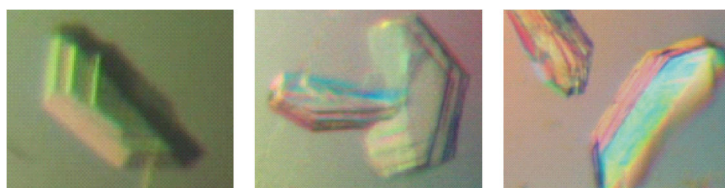
**Abbildung 4.15: 1.65 Å Kristallstruktur der Epimerisierungsdomäne aus der Tyrocidin Synthetase A**  
(A) Gesamtansicht mit den beiden Subdomänen in unterschiedlichen Blautönen (N-terminale Subdomäne: hellblau, C-terminale Subdomäne: dunkelblau). Die Seitenkette des katalytisch wichtigen Histidinrestes H146 ist in orange dargestellt. Die Sekundärstrukturelemente sind beschriftet. (B) Blick ins aktive Zentrum der TycA-Epimerisierungsdomäne, in dem mehrere wohldefinierte Wassermoleküle vorliegen. Die Elektronendichte ist konturiert bei 1.5  $\sigma$ .



### 4.3.2 Die Epimerisierungsdomäne der Tyrocidin Synthetase B

#### Stand der Forschung zu Beginn des Projekts

Durch frühere Arbeiten von Mohammad Mofid und Daniel Stein (Arbeitsgruppe Prof. Marahiel) war bereits eine Kristallisationsbedingung für die Epimerisierungsdomäne der Tyrocidin Synthetase B (TycB3-E) bekannt. Die Kristallisation im *hanging drop* bei 19 °C aus 2 M Na-Formiat, 0.1 M Na-Acetat lieferte verwachsene, hexagonal-scheibenförmige Kristalle (Abbildung 4.16).



**Abbildung 4.16: Verwachsene Kristalle der TycB3-Epimerisierungsdomäne**

Die Kristalle der Formiat/Acetat-Bedingung zeigen Fehlordnungen. Die Kristalle dieser Bedingung sind farblos und erscheinen wegen ihres Dichroismus nur bei Verwendung gekreuzter Polfilter farbig.

Die mangelhafte Fernordnung dieser Kristallform, die sich makroskopisch in einem unregelmäßigen Wachstum äußert, zeigt sich auch bei der kristallographischen Untersuchung in Form stark verschmierter Reflexe auf den Diffraktionsbildern. Die hohe Mosaizität dieser Kristallform ließ keine Indizierung und Integration der Reflexe und damit auch keine kristallographische Analyse zu.

#### Optimierung der Kristallisationsbedingung und Datensammlung

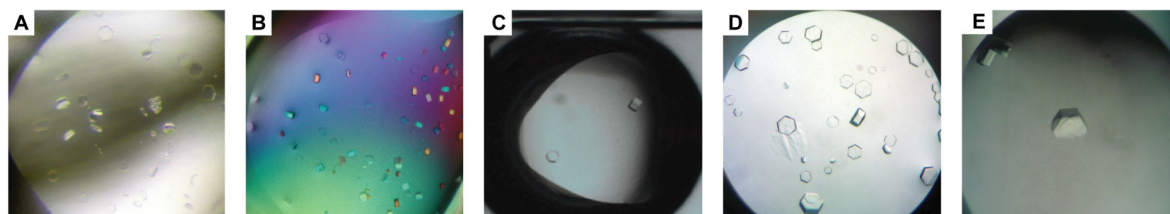
Am Max-Planck-Institut für Biophysik, Frankfurt/Main, wurden für die TycB3-Epimerisierungsdomäne Kristallisationsscreens mit Hilfe eines Kristallisationsroboters erstellt. Die Ansätze lieferten mit 0.05 M Na-Citrat (pH 5.6), 10 % Isopropanol und 10 % PEG 4000 flache, hexagonale Kristalle. Im Gegensatz zu den Kristallen der Acetat/Formiat-Bedingung wiesen diese an den Seiten entlang der sechszähligen Achse gleichmäßige Flächen und Kanten auf. Die so erhaltenen Kristalle aus einem Screen im 96-well-Maßstab sind in Abbildung 4.17A dargestellt.

Die anschließende manuelle Reproduktion und Optimierung dieser Kristalle in 24-well-Platten zeigte, dass sich die Kristalle unter dieser Bedingung nur in *sitting drop*-Ansätzen, nicht jedoch in *hanging drop*-Ansätzen, reproduzieren ließen. Außerdem bildeten sich meist viele kleine Kristalle, die für die röntgenkristallographische Untersuchung zu klein waren. Für die Kristallisation der TycB3-Epimerisierungsdomäne konnte keine Bedingung gefunden werden, unter der in Analogie zur PCP-C Bidomäne (vgl. Kapitel 4.2.2) reproduzierbar nur wenige große Kristalle erhalten werden. Repräsentative Kristalle dieser Optimierung sind in Abbildung 4.17B wiedergegeben.

## 4 Ergebnisse

Im Screen MBC II wurden ebenfalls äußerlich gute Kristalle in Tropfen aus 0.1 M Lithiumsulfat, 0.1 M Natriumcitrat (pH 5.6) und 12 % PEG 4000 beobachtet. Diese sind in Abbildung 4.17C dargestellt.

Die Kristallisationstropfen in Abbildung 4.17D und E zeigen, dass im Rahmen der Optimierung große, äußerlich gute Kristalle dieser Epimerisierungsdomäne erhalten wurden. Allerdings konnte die Bildung mehrerer Kristallisationskeime nicht immer unterdrückt werden.



**Abbildung 4.17: Kristalle der TycB3-Epimerisierungsdomäne aus citrathaltigen Kristallisationslösungen**  
(A) Erste Kristalle der neuen Bedingung aus 96-well-Screen. (B) Manuell reproduzierte Kristalle in 2 µL Tropfen (Aufnahme mit Polfilter). (C) Überraschend wenige, dicke Kristalle in einer Lithiumsulfat-haltigen Lösung des Screens MBC I. (D) und (E) Kristalle geeigneter Größe in Tropfen der Optimierung mit nativer bzw. Selenomethionin-substituierter Probe von TycB3-E.

### Aufnahme von Datensätzen

Die ersten Diffraktionsbilder eines Kristalls der TycB3-E-Domäne wurden mit MOSFLM indiziert. Demnach weist diese Kristallform eine sehr große asymmetrische Einheitszelle mit einer großen Kantenlänge entlang der hexagonalen Achse auf. In Anbetracht der ohnehin schon hohen Mosaizität dieser Kristallform wurden die Kristalle von da an so positioniert, dass die lange, hexagonale Achse bei Drehung um die  $\phi$ -Achse nicht parallel zum Röntgenstrahl liegt. Auf diese Weise sollte die gleichzeitige Aufnahme von dicht beieinanderliegenden Reflexen (hervorgerufen durch Beugung an den weit auseinanderliegenden Ebenen entlang der Achsen c) möglichst vermieden werden.

Es wurden zwei Datensätze von Kristallen der TycB3-E-Domäne aufgenommen. Bei dem ersten handelt es sich um einen Datensatz des nativen Proteins, beim zweiten um einen Datensatz des Selenomethionin-substituierten Proteins.

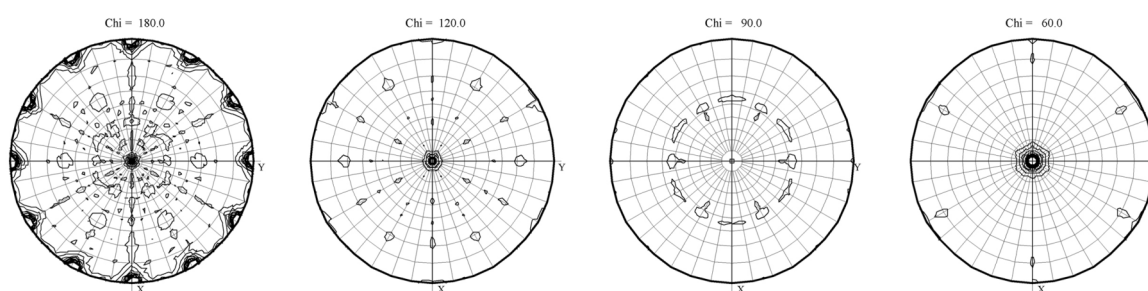
Der Datensatz des nativen Proteins wurde an der beamline BW7A des EMBL (DESY, Hamburg) aufgenommen. Aus der Indizierung ergab sich eine asymmetrische Einheitszelle mit den Kantenlängen  $a = 209.0 \text{ \AA}$ ,  $b = 121.0 \text{ \AA}$ ,  $c = 177.9 \text{ \AA}$  sowie mit den Winkeln  $\alpha = \gamma = 90.0^\circ$  und  $\beta = 94.7^\circ$ . Überraschenderweise ließ sich der Datensatz nur in der Raumgruppe C2 prozessieren. Damit lag entgegen der erwarteten sechszähligen nur eine zweizählige Symmetrieachse vor.

Allein mit diesen Daten und ohne Kenntnis der Phaseninformation ließ sich keine Elektronendichte generieren, um die Struktur selbst zu lösen. Angesichts der großen Einheitszelle ergab die Berechnung des Matthews-Koeffizienten keine eindeutige Lösung. Vielmehr erstreckte sich der



Bereich der wahrscheinlichen Molekülanzahl pro asymmetrische Einheitszelle von 6 bis 12, was Solvensgehalten von 67 % bis 34 % entsprach. Eine detaillierte Tabelle mit den Matthews-Koeffizienten in Abhängigkeit von der Molekülanzahl befindet sich im Anhang (Kapitel 8.6).

Da in der asymmetrischen Einheitszelle mehrere Proteinmoleküle erwartet wurden, sollte die Patterson-Selbstrotationsfunktion Auskunft über mögliche nichtkristallographische Symmetrien geben. Bei dieser Methode wird eine rotierte Version der Patterson-Funktion mit der Ausgangs-Patterson-Funktion überlagert. Kommen die beiden Funktionen dabei zur Deckung, nimmt die Selbstrotationsfunktion große Werte an. Die für den Datensatz der Epimerisierungsdomäne erhaltenen Werte sind in Abbildung 4.18 in Abhängigkeit der Rotationsparameter  $\theta$ ,  $\varphi$  und  $\chi$  dargestellt.



**Abbildung 4.18:** Ergebnis der Selbstrotationsfunktion für den nativen Datensatz von TycB3-E

Sämtliche Patterson-Selbstrotationsfunktionen wurden jeweils für die Winkel  $\chi = 180^\circ$ ,  $120^\circ$ ,  $90^\circ$  und  $60^\circ$  mit dem Programm MolRep [Vagin, 1997] erstellt.

Der Datensatz des Selenomethionin-substituierten Proteins TycB3-E war an der beamline ID 23-1 (ESRF, Grenoble) aufgenommen worden. Das derivatisierte Protein kristallisierte in derselben Kristallform. Es wurden vergleichbare Resultate hinsichtlich der Zellparameter, Raumgruppe und Selbstrotationsfunktion erhalten. Das Signal der anomalen Streuer (9 Selenatome pro Proteinmolekül) war jedoch nicht ausreichend. Es lieferte nicht die Phaseninformation, die zur Lösung der Struktur erforderlich wäre.

Weitere Datensätze wurden von Kristallen aufgenommen, die zuvor mit Schwermetallionen bzw. Schwermetallkomplexionen behandelt worden waren (*Soaking*). Da im Kristallisationspuffer das komplexbildende Citrat enthalten ist, erschien es wenig erfolgsversprechend, das *Soaking* mit Schwermetallkationen durchzuführen. Daher wurde versucht, das Protein vor der Kristallisation durch Behandlung mit  $\text{Hg}^{2+}$  an dem einzigen Cysteinrest, C449, kovalent zu modifizieren.

Daneben wurden Kristalle mit Wolframat und dem Komplexbkation  $\text{Ta}_6\text{Br}_{12}^{2+}$  behandelt. Auch diese Versuche der Derivatisierung lieferten keine Phaseninformationen. Deswegen hatten die nachfolgenden Versuche das Ziel, die Struktur mit Hilfe des *molecular replacements* zu lösen. Diese Versuche der Strukturlösung sind im Anhang zusammenfassend beschrieben (Kapitel 8.7).

### 4.3.3 Weitere bearbeitete Epimerisierungsdomänen

Die anderen Epimerisierungsdomänen, die in Abbildung 4.10 abgebildet sind, jedoch noch nicht weiter erwähnt wurden, sind die PCP-E-Bidomänen aus den Modulen TycB3, GrsA bzw. FenD2. Während mit den rekombinanten Proteinen TycA-E und FenD2-PCP-E keine Kristalle erhalten wurden, konnten mit dem Protein PCP-E aus TycB3 in einer Lösung aus 0.2 M Natriumacetat, 0.1 M Natriumcacodylat (pH 5.6) und 12.5 % PEG 8000 sehr dünne, verwachsene Kristalle beobachtet werden. Allerdings waren die Kristalle für die kristallographische Analyse zu klein und die Kristallqualität ließ sich nicht verbessern. Kleine Kristalle entstanden ebenfalls von GrsA-PCP-E unter zwei Bedingungen des Qiagen Classics Screens (0.2 M Ammoniumacetat, 0.1 M Natriumcitrat (pH 5.6), 30 % PEG 4000 bzw. 0.1 M Natriumcitrat (pH 5.6), 20 % Isopropanol, 20 % PEG 4000). Allerdings ließ sich die Kristallisation mit selbst angesetzten Lösungen nicht reproduzieren.

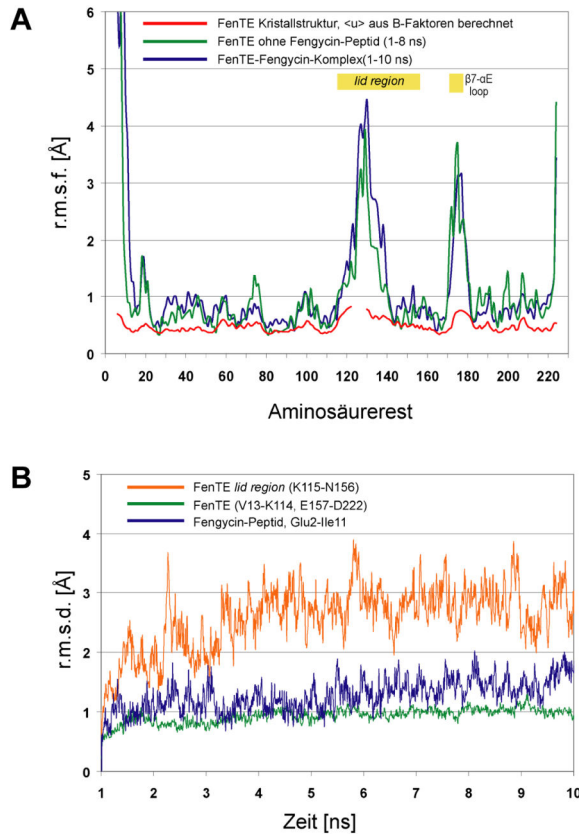
## 5 Diskussion

### 5.1 Das *edge-on* Modell des Fengycin Thioesterase-Fengycin-Komplexes

In der folgenden Diskussion zur Fengycin Thioesterase sollen der Bindungsmodus und die Wechselwirkungen im Enzym-Produkt-Komplex genauer betrachtet werden. Um eine Aussage über die biologische Relevanz der im Modell beobachteten Wechselwirkungen zu erhalten, wurde diese mit Hilfe biochemischer Methoden überprüft. Hierfür wurden Substratpeptide gezielt veränderter Sequenz synthetisiert. Dem Modell zufolge sollten sich einige der eingeführten Veränderungen auf die enzymatische Aktivität auswirken. Eine biochemische Überprüfung des Modells wurde von Björn Wagner (Arbeitsgruppe Prof. Marahiel) durchgeführt und ist hier nur zusammenfassend beschrieben. Daran schließt sich in Kapitel 5.2 eine Diskussion über die Substratbindung in TE-Domänen und ihre Interaktionen mit den davorliegenden PCP-Domänen an.

Die MD-Simulation des FenTE-Fengycin-Komplexes lief über 10 ns. Über den betrachteten Zeitraum war keine Dissoziation des Komplexes zu beobachten. Wie schon bei der vorbereitenden Simulation der FenTE ohne das Produktpeptid, wurde auch hier eine anfängliche Äquilibrierungsphase beobachtet. Danach blieben größere strukturelle Veränderungen aus.

Die Bindung des Fengycin Lipopeptids ändert qualitativ nichts an der Mobilitätsverteilung in der Thioesterase. Abbildung 5.1A stellt die Fluktuationen der einzelnen Reste einerseits in der Kristallstruktur und andererseits in den beiden MD-Simulationsläufen (mit und ohne gebundenes Fengycin-Peptid) dar. In beiden Simulationen zählen die *lid region* und der  $\beta 7$ - $\alpha$ E-Loop neben dem unnatürlichen N-Terminus zu den beweglichsten Bereichen des Enzyms. Die gemittelte Mobilität der *lid region* ist höher als die des restlichen Proteins, wie aus der separaten Betrachtung in Abbildung 5.1B hervorgeht. Dort sind die Abweichungen der Hauptkettenatome (nur  $C_{\alpha}$ , C und N) relativ zum äquilibrierten Zustand nach 1 ns aufgetragen. Abbildung 5.1C veranschaulicht, dass diese thermischen Bewegungen jedoch nicht die Bindung des Peptids an die Thioesterase beeinträchtigen: Der dem Solvens unzugängliche Teil der Oberfläche des Proteins und des Peptids bleibt während der gesamten Simulation praktisch unverändert.



**Abbildung 5.1: Analyse des FenTE-Fengycin-Komplexes**

(A) Fluktuationen der Aminosäurereste in der Kristallstruktur (rot) und den MD-Simulationen mit (blau) und ohne Fengycin-Peptid (grün).

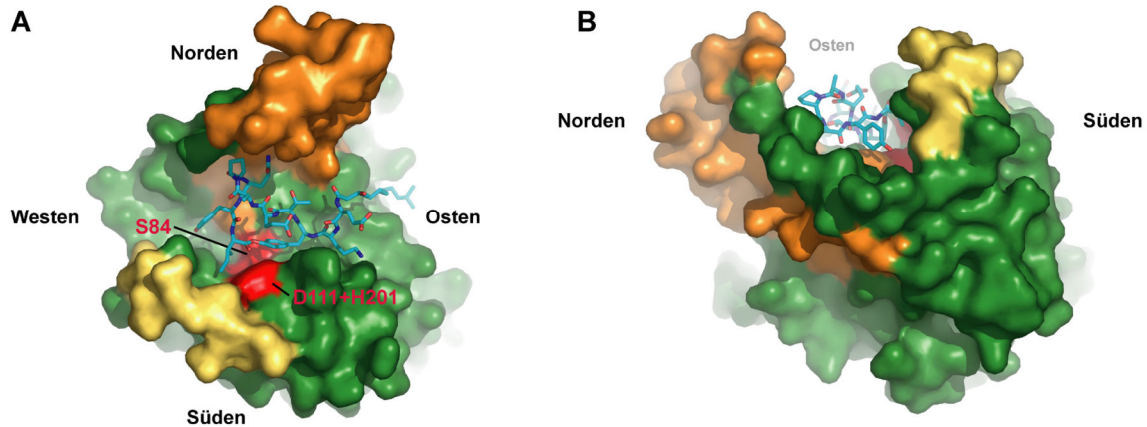
(B) Gemittelte Abweichungen der Hauptkettenatome ( $C_\alpha$ , C, N) der FenTE lid region, des FenTE Kernproteins sowie des Produktpeptids relativ zu ihren Positionen nach einer 1 ns.

(C) Oberflächen der Thioesterase, des Fengycin-Peptids, der Kontaktoberfläche und des Komplexes in Abhängigkeit von der Zeit.

Einen Überblick über die Bindung des Lipopeptidprodukts an die Fengycin Thioesterase vermittelt Abbildung 5.2. Die Richtungsbezeichnungen (N, S, W, O) entsprechen denen, die Bruner *et al.* bei der Beschreibung der Surfactin TE benutzten [Bruner, 2002]. Aus der Aufsicht in Teilabbildung A ist zu erkennen, dass die Peptidbindungstasche in Ostrichtung nicht verschlossen ist. Da das Peptid somit auf der Kante der Bindungstasche aufsitzt, wurde der Bindungsmodus als *edge-on*-Bindung bezeichnet.

Aus der West-Perspektive, die in Teilabbildung B wiedergegeben ist, lässt sich die *Canyon*-ähnliche Form der Substratbindungstasche erkennen. Sie wird auf der Nordseite durch die lid region (orange) und auf der Südseite durch die  $\beta$ 7- $\alpha$ E- und  $\beta$ 8- $\alpha$ F-Loops (goldfarben bzw. grün) begrenzt.

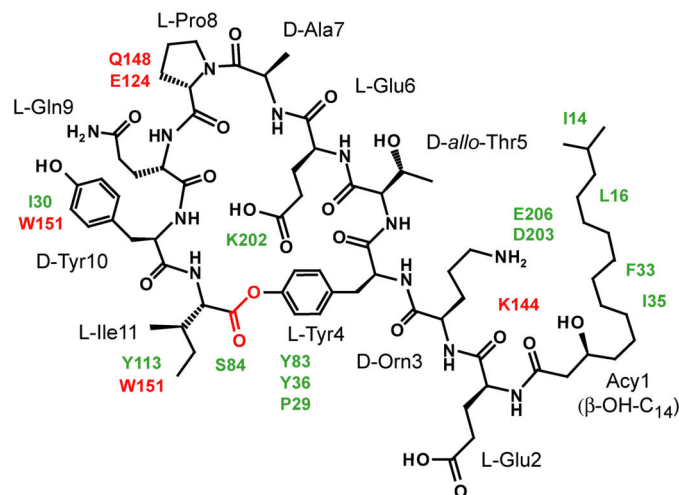
Aus den Daten der Simulation lassen sich sämtliche Wechselwirkungen entnehmen. Die Analyse der nicht-kovalenten Wechselwirkungen, d. h. Wasserstoffbrückenbindungen sowie ionische und *van der Waals*-Wechselwirkungen, erfolgte durch Bestimmung der Atomabstände und ihrer zeitlichen Änderungen.



**Abbildung 5.2: Komplexstruktur der Fengycin Thioesterase mit dem Produktpeptid Fengycin**

**(A)** Aufsicht auf die Thioesterasedomäne FenTE. Die Reste der katalytischen Triade sind rot, die *lid region* ist orange und der  $\beta 7$ - $\alpha$ E-Loop ist goldfarben hervorgehoben. **(B)** Ansicht des Komplexes von der Westseite: Aus dieser Perspektive wird die *Canyon*-ähnliche Form der Bindungstasche deutlich.

In biochemischen Untersuchungen mit den Thioesterasedomänen der Surfactin und Tyrocidin Biosynthesecuster ließ sich zeigen, dass die Enzyme nur mit wenigen Resten des jeweiligen Substratpeptids in Kontakt treten [Trauger, 2000; Tseng, 2002]. Diese sind die an der Zyklisierung direkt beteiligten Aminosäurereste des Substratpeptids und ggf. deren unmittelbar benachbarten Reste. Die Thioesterasedomänen zeigen dabei eine hohe Spezifität für deren Identität und stereochemische Konfiguration. Daher gilt ein besonderes Interesse denjenigen Aminosäureresten, die sich in unmittelbarer Nähe zu den zu verknüpfenden Resten, L-Tyr 4 und L-Ile 11, befinden. In Abbildung 5.3 sind die Aminosäurereste der Fengycin Thioesterase beschrieben, die mit dem Fengycin Produktpeptid interagieren.



**Abbildung 5.3: An der Substratbindung beteiligte Aminosäurereste der Fengycin Thioesterase**

Abbildung 105: An der Substratbindung beteiligte Aminosäurereste der Fengycin-Aminosterase. Die Reste des Fengycin-Peptids sind im Dreibuchstaben-Code (schwarz), die der Fengycin Thioesterasedomäne im Einbuchstaben-Code (grün) angegeben. Reste der *lid region* sind rot hervorgehoben.

Die Seitenketten der Aminosäurereste D-Tyr 10 und L-Ile 11 liegen in der äquilibrierten Struktur in hydrophoben Taschen. Die seitens des Proteins beteiligten Aminosäurereste sind I30, Y113 und W151. Der benachbarte Rest L-Glu 9 liegt auf der Nord-Seite des *Canyons* und interagiert dort mit den Resten der *lid region*, Q148 und E124.

Auf der Westseite des *Canyons* ist L-Tyr 4 in einer hydrophoben Spalte gebunden, die durch die Reste P29, Y36 und Y83 gebildet wird. Diese Bindung bringt den Rest L-Tyr 4 in die richtige Position für den nukleophilen Angriff auf den thioestergebundenen C-Terminus des Restes L-Ile 11.

Einige der beobachteten Wechselwirkungen bestehen jedoch nicht über den gesamten Simulationszeitraum. So interagiert beispielsweise der dem intramolekularen Nukleophil L-Tyr 4 benachbarte Aminosäurerest D-Orn 3 des Fengycinpeptids mit den zwei Seitenketten D203 und E206 der Fengycin Thioesterase. Während die Wechselwirkung mit dem Glutamatrest stabil ist, löst sich die Bindung zur Seitenkette des Restes D203 zunehmend nach etwa 5 ns. Dennoch ist die Wechselwirkung der Ornithinseitenkette mit den Gruppen des Enzyms für die Bindung des Peptids von Bedeutung. Dies zeigt sich in der Abnahme der enzymatischen Aktivität gegenüber einem D-Orn 3→D-Ala 3 substituierten Substratpeptid. Die Einführung eines sauren Restes an dieser Position erhöhte die Hydrolyserate um den Faktor 2, so dass diese Interaktionen offenbar die Positionierung des Substrats für die Zyklisierung beeinflussen.

Ebenso lagert sich der N-terminal gebundene Acylrest teilweise an eine hydrophobe Oberfläche an, die auf der Ostseite des *Canyons* von den Seitenketten der Aminosäuren I14, L16, F33 und I35 gebildet wird. Diese Bindung wird allerdings im Verlauf der Simulation gelöst. Offensichtlich ist die Wechselwirkung der Thioesterase mit der N-terminal gebundenen Fettsäure für die Zyklisierung nicht essentiell. Diese Beobachtung geht einher mit den Ergebnissen der nachfolgend beschriebenen biochemischen Untersuchungen, nach denen die Verkürzung der Fettsäure von C14 auf C4 keine merkliche Änderung in der enzymatischen Aktivität bewirkte.

Bedingt durch die *edge-on* Bindung an die Proteinoberfläche und die Konformation des Peptids an sich kommen die mittleren Reste, D-*allo*-Thr 5 bis L-Pro 8, kaum mit dem Enzym in Kontakt. Dennoch hatte die Substitution von L-Glu 6 durch L-Ala 6 einen merklichen Effekt auf den  $K_M$ -Wert (etwa Faktor 6) und den  $k_{cat}$ -Wert (etwa Faktor  $\frac{1}{2}$ ) der Zyklisierungsreaktion. Dies könnte allerdings auch auf einem Wegfallen intramolekularer Wasserstoffbrückenbindungen beruhen. Die Seitenkette des Glutamats interagiert nämlich zumindest zeitweise mit der Seitenkette von D-*allo*-Thr 5 und der Hauptkette von L-Tyr 4.

### Biochemische Überprüfung des Komplexmodells

Die Versuche zur Analyse der enzymatischen Aktivität wurden von Dr. Björn Wagner (Arbeitsgruppe Prof. Marahiel) durchgeführt. Dazu wurden Substratpeptide hergestellt, in denen an jeweils einer Position die Aminosäure verändert war. Diese Variationen in der Sequenz basierten auf dem Modell des Enzym-Produkt-Komplexes und sollten die produktive Wechselwirkung zwischen Protein und Peptid näher untersuchen. Die Substratpeptide wurden durch Kupplung mit Thiophenol aktiviert und mit dem rekombinanten Protein Fen-PCP-TE inkubiert. Durch anschließende Untersuchung der Reaktionsansätze mittels HPLC und MALDI-TOF bzw. ESI-MS, wurde die enzymatische Aktivität der Fengycin Thioesterase gegenüber den veränderten Substraten ermittelt. Die vorgenommenen Variationen beeinflussten die Substratakzeptanz in der vorhergesagten Weise und bekräftigten dadurch die Richtigkeit des Modells.

**Tabelle 5: Ergebnisse der kinetischen Untersuchung der Fengycin PCP-TE-katalysierten Zyklisierungsreaktion**

Substratpeptid	$K_M$ [mM]	$k_{cat}$ [min <sup>-1</sup> ]	$k_{cat} / K_M$ [min <sup>-1</sup> mM <sup>-1</sup> ]	Zyklisierung/ Hydrolyse	Substrat/ Produkt
FLP *	0.52	1.71	3.28	4.3	8.1
FLP-D-Orn3/D-Ala3	6.24	0.32	0.05	4.5	68.4
FLP-D-Orn3/D-Glu3	3.43	0.36	0.10	2.4	46.9
FLP-L-Glu6/L-Ala6	2.97	0.94	0.32	5.0	17.4
FLP-D-Ala7/D-Ser7	0.48	1.98	4.12	3.7	6.6
FLP-L-Gln9/L-Ala9	4.54	0.88	0.19	5.9	29.9
FLP-C14	0.45	2.02	4.49	3.5	7.2

Es wurden jeweils 200  $\mu$ M Substrat und 5  $\mu$ M Fen-PCP-TE für 30 min inkubiert.

\* Die Sequenz des synthetischen FLP lautet: (3-(R)-hydroxy-butyrat)1 – L-Glu2 – D-Orn3 – L-Tyr4 – D-Thr5 – L-Glu6 – D-Ala7 – L-Pro8 – L-Gln9 – D-Tyr10 – L-Ile11

Wie aus den kinetischen Daten der FenTE hervorgeht, wirkten sich die Substitutionen der Reste D-Orn 3, L-Glu 6 und L-Gln 9 jeweils nachteilig auf die enzymatische Aktivität der FenTE aus. Dem Modell des FenTE-Fengycin-Komplexes zufolge interagieren diese drei Reste mit dem Enzym. Daher stehen die Resultate der biochemischen Überprüfung in Einklang mit den im Komplexmodell vorgefundenen intramolekularen Wechselwirkungen.

Eine weitere Substitution sollte die Lage des Alaninrestes L-Ala 7 testen. Die Seitenkette des Restes L-Ala 7 ist dem Enzym abgewandt und geht keine Wechselwirkungen damit ein. Daher entspricht die Beobachtung, dass die Substitution keine nennenswerte Auswirkung auf die Aktivität hat, dem anhand des Modells erwarteten Verhalten.

## 5.2 Faktoren der Substratbindung in NRPS-Thioesterasedomänen

Das aktive Zentrum der Fengycin Thioesterasedomäne liegt in einer grabenförmigen Vertiefung, die einerseits von der flexiblen *lid region*, andererseits von den  $\beta 7$ - $\alpha E$ - und  $\beta 8$ - $\alpha F$ -Loops begrenzt wird. Die katalytische Triade, die das aktive Zentrum der Thioesterasedomäne markiert, liegt am Boden dieser Vertiefung.

Aus der Kristallstruktur der Surfactin Thioesterasedomäne (SrfTE) [Bruner, 2002] ist ersichtlich, dass auch in diesem Protein das aktive Zentrum am Boden einer Vertiefung liegt. Bruner *et al.* hatten ein Modell für die Zyklisierung vorgeschlagen, wonach die Konformation des zu zyklisierenden Surfactinpeptids hauptsächlich durch die Form dieser Vertiefung bestimmt würde. Das intramolekulare Nukleophil, die 3-Hydroxyfunktion der Fettsäure, kommt dadurch in räumliche Nähe des C-terminalen Restes L-Leu 7. Diese Positionierung ermöglicht den nukleophilen Angriff auf das Carboxyl-Kohlenstoffatom des Thioesters, was dann zur Zyklisierung des Substratpeptids führt.

Weiterführende Untersuchungen [Tseng, 2002] mit der Surfactin TE und substratähnlichen Inhibitoren zielten darauf ab, das Peptid kovalent im aktiven Zentrum zu binden, um dessen Konformation und seine Wechselwirkungen mit der Thioesterase kristallographisch zu untersuchen. Die Thioesterase wurde dazu mit einem Peptidyl-Borsäureester-Inhibitor umgesetzt, der kovalent an den katalytisch aktiven Serinrest der Thioesterase bindet. Allerdings ließen sich in der Elektronendichtekarte nur die seringebundene B(OH)-Gruppe sowie die beiden C-terminalen Reste D-Leu 6 und L-Leu 7 erkennen. Durch Überlagerung der NMR-Struktur des Lipopeptids Surfactin wurde ein Bindungsmodell abgeleitet, in dem die mittleren Reste des Peptids, d. h. die zum Nukleophil und zum Thioester entfernt gelegenen Reste, aus der Bindungstasche herausragen und keinen Kontakt zum Enzym haben.

Dieses Modell ist mit den zuvor an der Tyrocidin Thioesterase gewonnenen Erkenntnissen gut vereinbar. Es ließ sich zeigen, dass die Tyrocidin Thioesterase nur die Identität und die Konfiguration der Aminosäurereste in der Umgebung der an der Zyklisierung beteiligten Reste erkennt [Trauger, 2000]. Veränderungen im mittleren Bereich der Tyrocidinpeptids hingegen zeigten keine größeren Auswirkungen auf das Zyklisierungsverhalten der Thioesterase. Des Weiteren ließen sich mit der Thioesterase auch deutlich längere Substratpeptide zyklisieren, so dass Makrolaktame mit Ringgrößen von bis zu 14 Aminosäureresten zugänglich waren [Kohli, 2001]. Mit Hilfe des SrfTE Bindungsmodells lassen sich sowohl die fehlende Substratspezifität, als auch die Erweiterbarkeit zu größeren Substraten erklären. Dieses Modell erlaubt sowohl die Substitution entfernt gelegener Aminosäuren als auch eine Erweiterung des Makrozyklus, da die Substratbindung nicht durch die Größe der Substratbindungstasche limitiert wird.



Es wird davon ausgegangen, dass die enzymatische Makrozyklisierung durch die Vorfaltung (*pre-organization*) des Substratpeptids unterstützt wird. Die strukturelle Untersuchung der Makrolaktamverbindungen Gramacidin S und Tyrocidin ergab, dass die Peptide die Konformation eines antiparallelen  $\beta$ -Faltblatts annehmen, die durch intramolekulare Wasserstoffbrückenbindungen stabilisiert wird [Kuo, 1980]. Es ist daher zu vermuten, dass sich eine solche Struktur bereits in den linearen Vorläuferpeptiden ausbildet. Auf diese Weise kämen die an der Zyklisierung unmittelbar beteiligten Reste in räumliche Nähe zueinander. Demnach wäre der Einfluss der Substratbindungstasche bei der Zyklisierung dieser vorgefalteten Peptide weniger entscheidend. Es wurde sogar beschrieben, dass aktivierte, lineare Vorläuferpeptide, in denen eine solche Vorfaltung vorliegt, auch in Abwesenheit einer Thioesterase zyklisieren [Bu, 2002].

Im Gegensatz zur Tyrocidin Thioesterase sind die Thioesterasen der Surfactin und Fengycin Synthetasecluster nicht so tolerant gegenüber Veränderungen im Substratpeptid. Zwar gelang mit der Surfactin TE durch Verlagerung der Fettsäure-Hydroxylgruppe an die  $\epsilon$ -Position eine Erweiterung der Ringgröße um 3 Atome, mit einer  $\alpha$ -positionierten Hydroxylgruppe hingegen schlug die Zyklisierung zu einem um ein Atom verkleinerten Lakton fehl [Tseng, 2002]. Die Fengycin TE katalysierte zwar ebenfalls die Zyklisierung eines Substrats zu einem um einen Aminosäurerest vergrößerten Ring (Ringgröße 32 Atome anstelle von 29), die Reaktion zu einem verkleinerten Lakton wurde aber auch nicht katalysiert. Statt einer Zyklisierung zu einem 26 Atome enthaltenden Zyklus wurde ausschließlich die Hydrolyse dieses Substrats beobachtet [Sieber, 2003].

Die Makrolaktone Fengycin und Surfactin enthalten im Gegensatz zu Gramacidin S und Tyrocidin keine  $\beta$ -Faltblatt-Struktur oder eine vergleichbar ausgeprägte, intramolekular stabilisierende Struktur. Daher wird in den linearen Vorläuferpeptiden die Vorfaltung zu einem geringeren Maß vorliegen. Es ist zu vermuten, dass in diesen Fällen die Thioesterasedomänen durch ihre Wechselwirkungen mit den Substratpeptiden einen größeren Einfluss auf deren Konformation ausüben. Die resultierende größere Abhängigkeit von den Enzym-Substrat-Interaktionen könnte die beobachteten unterschiedlichen Verhalten in den Zyklisierungsreaktionen erklären.

Das in der vorliegenden Arbeit vorgestellte Bindungsmodell eines FenTE-Fengycin-Komplexes beschreibt einen *edge-on* Bindungsmodus des Substrats an einer *Canyon*-ähnlichen Protein-oberfläche. Das Modell lässt einerseits intermolekulare Wechselwirkungen an der gemeinsamen Kontaktfläche zu, andererseits erklärt es, warum manche Reste, obwohl sie in der Nähe des aktiven Zentrums liegen, an der Substraterkennung und Substratfaltung nicht beteiligt sind. Eine Variation dieser Reste sollte die Akzeptanz des Substrats nicht beeinflussen.

### Der Einfluss des intramolekularen Nukleophils

Die Ergebnisse der von Sieber *et al.* vorgenommenen Untersuchungen an der Fengycin Thioesterase [Sieber, 2003] lassen sich mit denen der hier beschriebenen, strukturellen Arbeiten vergleichen. Sieber *et al.* untersuchten die Chemo- und Regioselektivität der Thioesterasedomäne, indem sie die enzymatische Aktivität gegenüber veränderten Substratpeptiden analysierten. Dazu wollten sie mit Hilfe einer L-Tyr 4→L-Phe 4-Mutation überprüfen, ob die nukleophilen Seitenketten der benachbarten Aminosäuren, D-Orn 3 und D-Thr 5, den Tyrosinrest als intramolekulares Nukleophil ersetzen können. Dieses Substrat wurde aber ausschließlich hydrolysiert, was dafür spricht, dass keiner der beiden benachbarten Reste als Nukleophil wirkte.

Um zu klären, ob sterische Gründe die Zyklisierung verhinderten oder die Thioesterase für die Tyrosinseitenkette spezifisch ist, wurde das Nukleophil Tyrosin gegen die benachbarten Aminosäuren, D-Orn 3 und D-Thr 5 ausgetauscht. Die Umsetzung des Substrats mit D-Tyr 3 und L-Orn 4 ergab neben dem linearen Hydrolyseprodukt auch ein zyklisches Produkt mit einer vergrößerten Ringgröße. Dagegen wurde mit dem die Aminosäuren L-Thr 4 und D-Tyr 5 enthaltenden Substrat lediglich das lineare Produkt der Hydrolyse erhalten.

Insgesamt zeigen diese Ergebnisse, dass die Zyklisierung ausschließlich über den Tyrosinrest erfolgt. Auch die Positionierung der Aminosäuren Ornithin und Threonin auf die natürliche Position des Nukleophils führte nicht zur Zyklisierung. Im ersten Fall wurde sogar das um eine Position verschobene Tyrosin dem Ornithin vorgezogen.

Für das Ausbleiben der Zyklisierung zu einem verkleinerten Laktone können sterische Gründe verantwortlich sein. Die Umpositionierung könnte bewirken, dass der Rest L-Tyr 5 in dem für die Zyklisierung gefalteten Peptid seine Nukleophil-Bindungsstelle nicht mehr erreicht.

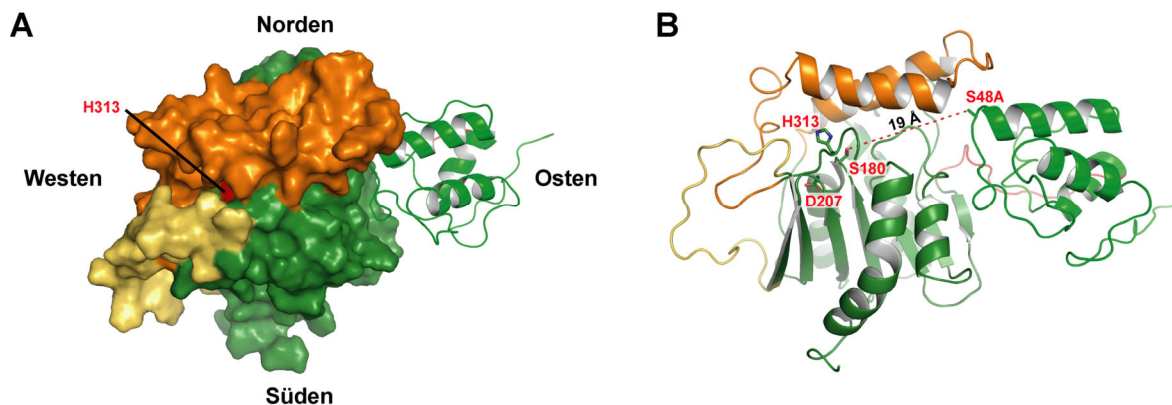
Bezogen auf das Bindungsmodell bedeutet dies, dass die Bindungstasche, welche in der Komplexstruktur die Tyrosinseitenkette bindet, für diese Seitenkette spezifisch ist. Die Möglichkeit der Laktoneerweiterung spricht für einen Bindungsmodus, in dem die Substratbindungstasche nicht limitierend wirkt. Diese beiden Punkte lassen sich mit dem *edge-on* Modell vereinbaren.

### Wechselwirkungen zwischen PCP- und TE-Domänen

Kürzlich wurde über PCP-TE-Interaktionen berichtet, die aus einer NMR-Studie erhalten wurden. Es handelt sich dabei um die PCP-TE-Bidomäne aus EntF, einer Synthetase des Enterobactin Biosynthesecusters aus *E. coli* (PDB: 2ROQ) [Frueh, 2008]. In dem untersuchten Enzym ist der katalytisch aktive Serinrest zu Alanin mutiert. Damit ist weder das Substrat, noch der Kofaktor an das Enzym gebunden, so dass sich nur die Interaktionen zwischen den Domänen und ihre Konformationen untersuchen ließen. In der dort beobachteten Anordnung der beiden Domänen sind

die aktiven Zentren 19 Å voneinander entfernt (Abbildung 5.4). Da dies eine produktive Substratübertragung auf die Thioesterasedomäne gestattet, kann es sich hierbei um eine biologisch relevante Anordnung handeln. Neben den Interaktionen zwischen den Domänen, die sich erstmals analysieren ließen, zeigt die Struktur auch, dass sich die PCP-Domäne an der Ostseite der Thioesterase anlagert. Bruner *et al.* hatten in der Kristallstruktur der Surfactin TE ein an das Enzym angelagertes Sulfation beobachtet [Bruner, 2002]. Sie nahmen an, dass im natürlichen *holo*-Protein diese Bindungstasche die Phosphatgruppe des 4'-Phosphopantethein-Kofaktors bindet. Sie schlugen ein Modell für die PCP-TE-Interaktion vor, bei dem sich die PCP-Domäne an die Ostseite der TE-Domäne anlagert. Vergleicht man dieses Modell mit der Struktur der EntF-PCP-TE-Bidomäne, so zeigt sich, abgesehen von geringfügigen Unterschieden hinsichtlich der Orientierung der PCP-Domäne und damit der Position des katalytisch aktiven Serinrestes, eine prinzipielle Übereinstimmung. Diese Unterschiede können darin begründet sein, dass sich die Domänenanordnung in den Enzymen des Enterobactin Biosyntheseclusters von der des Surfactin Biosyntheseclusters unterscheidet.

Das in dieser Arbeit entwickelte Komplexmodell der Fengycin Thioesterase kommt zu einer ähnlichen Vorhersage. Die *Canyon*-begrenzenden Bereiche, nämlich die *lid region* und der  $\beta$ 7- $\alpha$ E-Loop, verschließen den Zugang zum *Canyon* über die Westseite. Da das aktive Zentrum zudem weder über die Nord- noch über die Südseite zugänglich ist, ist anzunehmen, dass die Interaktion mit der PCP-Domäne auf der Ostseite der Thioesterase erfolgt.



**Abbildung 5.4: NMR-Struktur der PCP-TE-Bidomäne aus dem Enterobactin Biosynthesecluster**

Die PCP-Domäne und die Thioesterasedomäne sind grün dargestellt. Die *bridge region* ist orange, der  $\beta$ 7- $\alpha$ E-Loop goldfarben hervorgehoben. Die PCP-Domäne ist jeweils in der Cartoon-Repräsentation dargestellt.

(A) Die PCP-Domäne liegt auf der Ostseite der Thioesterasedomäne. Die *lid region* verdeckt die Substratbindungstasche, so dass von der katalytischen Triade ausschließlich H313 sichtbar ist. (B) Ansicht der PCP-TE-Bidomäne von Süden. Die Aminosäurereste A48 und S180 liegen etwa 19 Å voneinander entfernt. Die beobachtete relative Orientierung ließe eine Wechselwirkung zwischen den beiden Domänen zu.

### 5.3 Die Struktur der PCP-C-Bidomäne

Im Rahmen der vorliegenden Arbeit wurde die Struktur des Bidomänenproteins PCP-C bis zu einer Auflösung von 1.85 Å bestimmt. Das rekombinante Protein umfasst eine PCP-Domäne und eine C-Domäne aus der Synthetase TycC des Tyrocidin Biosyntheseclusters aus *B. brevis* ATCC 8185 (GenBank Eintrag: O30409).

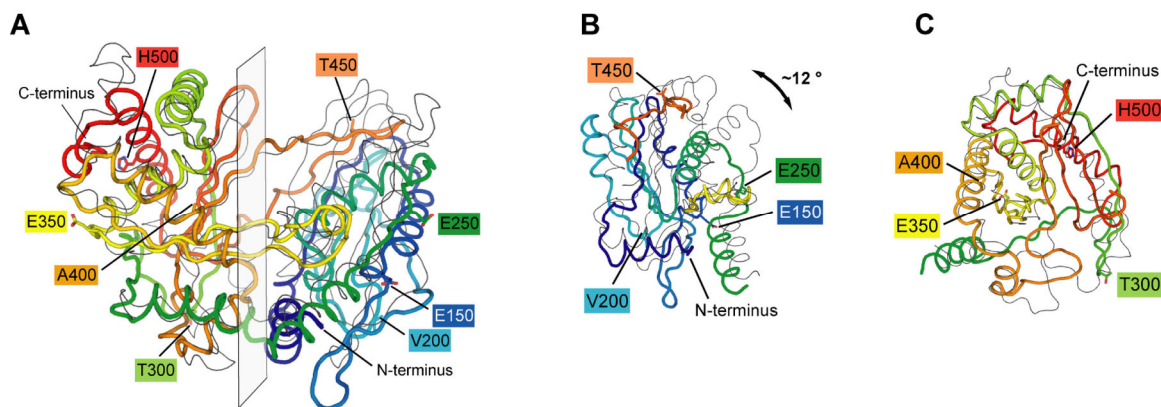
In der Kristallstruktur des 540 Aminosäuren großen, rekombinanten Proteins sind die Aminosäurereste R3–M455 und von Q461-L522 in der Elektronendichtekarte definiert. Die Termini sowie die dazwischenliegenden Reste T456-G460 sind hingegen nicht erkennbar. In der Kristallstruktur des Proteins lassen sich zwei Domänen erkennen, und zwar eine PCP-Domäne (M1-T82) und eine deutlich größere Kondensationsdomäne (V101-L522). Beide Domänen sind über einen sogenannten *Linker* (A83-P100) miteinander verbunden. Durch die Kristallisation des Bidomänenproteins im Rahmen der vorliegenden Arbeit war es erstmals möglich, den Verlauf des dazwischenliegenden Bereichs und dessen Wechselwirkungen mit den beiden Domänen zu untersuchen.

Die PCP-Domäne bildet eine abgeschlossene Einheit und setzt sich aus 4  $\alpha$ -Helices zusammen. Diese Anordnung, die bei mehreren *Carrier*-Proteinen des primären und sekundären Metabolismus beobachtet wird, wird als „4-Helix-Bündel“ bezeichnet [Weber, 2000]. Der strukturelle Vergleich der PCP-Domäne aus PCP-C mit den von Koglin *et al.* beschriebenen Zuständen A, A/H und H (PDB: 2GDY, 2GDW und 2GDX) ergab, dass die PCP-Domäne in PCP-C eine A/H-ähnliche Konformation aufweist [Koglin, 2006]. Die Überlagerung mit der A/H-ähnlichen Struktur von Weber *et al.* ergab einen r.m.s.d.-Wert von 1.67 Å (für 56 C $_{\alpha}$ ). Die analogen Überlagerungen mit den A- bzw. H-Zuständen offenbarten stärkere Abweichungen, was sich in größeren r.m.s.d.-Werten widerspiegelte.

Die Kondensationsdomäne aus PCP-C besteht aus zwei Subdomänen. Sie ähnelt damit VibH, der bis dahin einzigen, strukturell charakterisierten NRPS-Kondensationsdomäne. Die beiden Subdomänen, die als N- und C-terminale Subdomäne bezeichnet werden [Keating, 2002], sind V-förmig angeordnet. Jede dieser beiden Subdomänen zeigt strukturelle Ähnlichkeit mit den Monomeren der Chloramphenicol-Acetyltransferase CAT [Lewendon, 1994] und mit der Dihydro-lipoamid-Transacetylase E2p der bakteriellen Pyruvat Decarboxylase [De Crecy-Lagard, 1995]. CAT liegt in ihrer katalytisch aktiven Form als Homotrimer vor. In diesem Trimer werden drei aktive Zentren an den Grenzflächen zweier Monomere ausgebildet. In jedem aktiven Zentrum sind Aminosäurereste beider Monomere an der Katalyse beteiligt. In vergleichbarer Weise liegen die katalytisch wichtigen Aminosäurereste und damit vermutlich das aktive Zentrum der TycC6-

C-Domäne an der Grenzfläche der beiden Subdomänen. Die Struktur der Kondensationsdomäne kann somit als Pseudodimer des CAT-Faltungstyps angesehen werden.

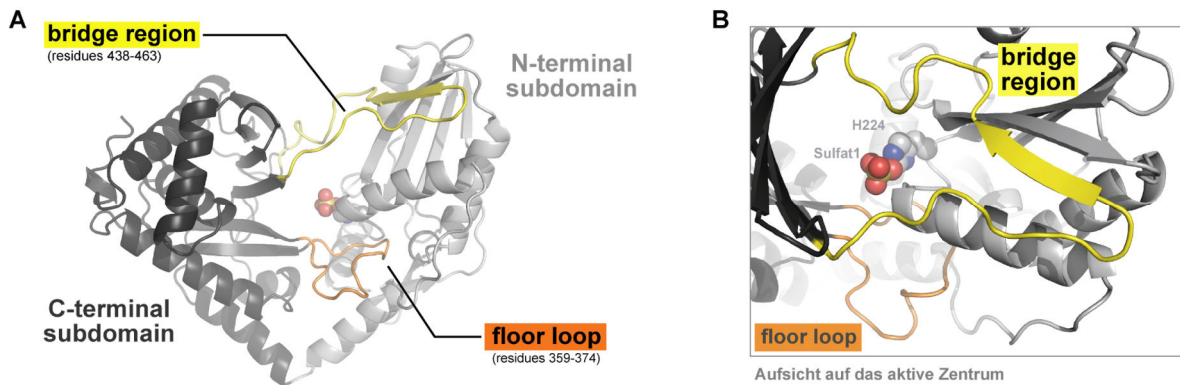
Trotz der mäßigen Sequenzidentität von nur 19 % zeigen die C-Domäne aus PCP-C und VibH (PDB: 1L5A) eine ausgeprägte strukturelle Ähnlichkeit. Die Überlagerung beider Proteine, die in Abbildung 5.5A dargestellt ist, macht dies deutlich und ergibt einen r.m.s.d.-Wert von 1.58 Å (für 197 C<sub>α</sub>). Allerdings fällt auf, dass dabei nur die C-terminalen Subdomänen aufeinander zum Liegen kommen, während die N-terminalen Subdomänen gegeneinander verschoben erscheinen. Tatsächlich weichen die relativen Orientierungen der beiden Subdomänen in PCP-C und VibH um 12 ° voneinander ab (Abbildung 5.5B und C). Das Zentrum dieser Verdrehung der beiden Subdomänen liegt an deren Übergang zwischen den beiden Helices α5 und α6 (PCP-C: S268; VibH: S174). Werden die beiden vergleichbaren Subdomänen separat überlagert, so findet man bei den N-terminalen Subdomänen (r.m.s.d. 1.70 Å für 116 C<sub>α</sub>) eine etwas schlechtere Übereinstimmung als bei den C-terminalen Subdomänen (r.m.s.d. 1.35 Å für 149 C<sub>α</sub>).



**Abbildung 5.5: Überlagerung der Strukturen der TycC6-C-Domäne und VibH**

(A) Die Struktur von VibH ist als dünne, schwarze Linie dargestellt. Die C-Domäne aus PCP-C ist farbig dargestellt, wobei die Farbe von blau (N-Terminus) nach rot (C-Terminus) wechselt. Zur besseren Vergleichbarkeit mit den Teilabbildungen (B) und (C) wurde jeder 50. Rest markiert. (B) und (C) Querschnitt-Ansicht der beiden Subdomänen nach Teilung entlang der in (A) angedeuteten Schnittfläche.

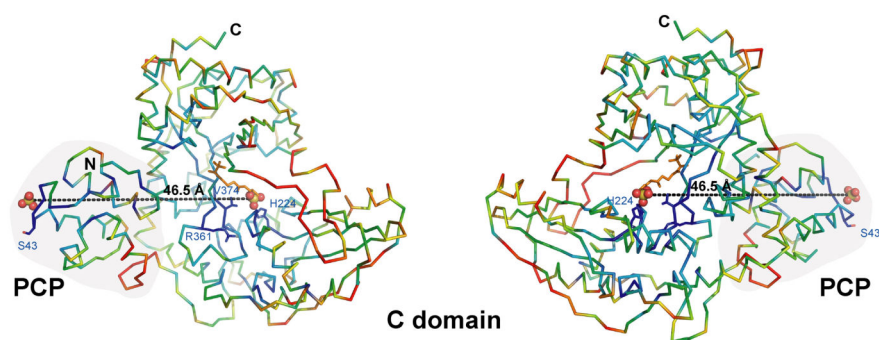
Zwischen den beiden Subdomänen gibt es lediglich zwei Bereiche, in denen sie miteinander in Wechselwirkung treten. Einer dieser intramolekularen Kontaktbereiche bildet den Boden des aktiven Zentrums und wurde als *floor loop* bezeichnet. Der bogenförmige Verlauf setzt sich aus den zwischen den β-Strängen β8 und β9 liegenden Resten T359-V374 zusammen, die in Abbildung 5.6 orange hervorgehoben sind. Den anderen Kontaktbereich stellt die *bridge region* dar (gelb in Abbildung 5.6), die das aktive Zentrum überbrückt und die Reste N438-A463 einschließt. In beiden Fällen handelt es sich um Bereiche der C-terminalen Subdomäne, die mit der N-terminalen Subdomäne in Kontakt treten.



**Abbildung 5.6: Farbliche Hervorhebung der zwei Kontaktbereiche zwischen N- und C-terminaler Subdomäne**  
**(A)** Die als *bridge region* und *floor loop* bezeichneten Ausläufer der C-terminalen Subdomäne bilden die einzigen nennenswerten Wechselwirkungen zwischen den beiden CAT-ähnlichen Subdomänen. **(B)** Beide liegen in nächster Umgebung des aktiven Zentrums, dessen Lage durch den Histidinrest H224 und das gebundene Sulfation markiert wird.

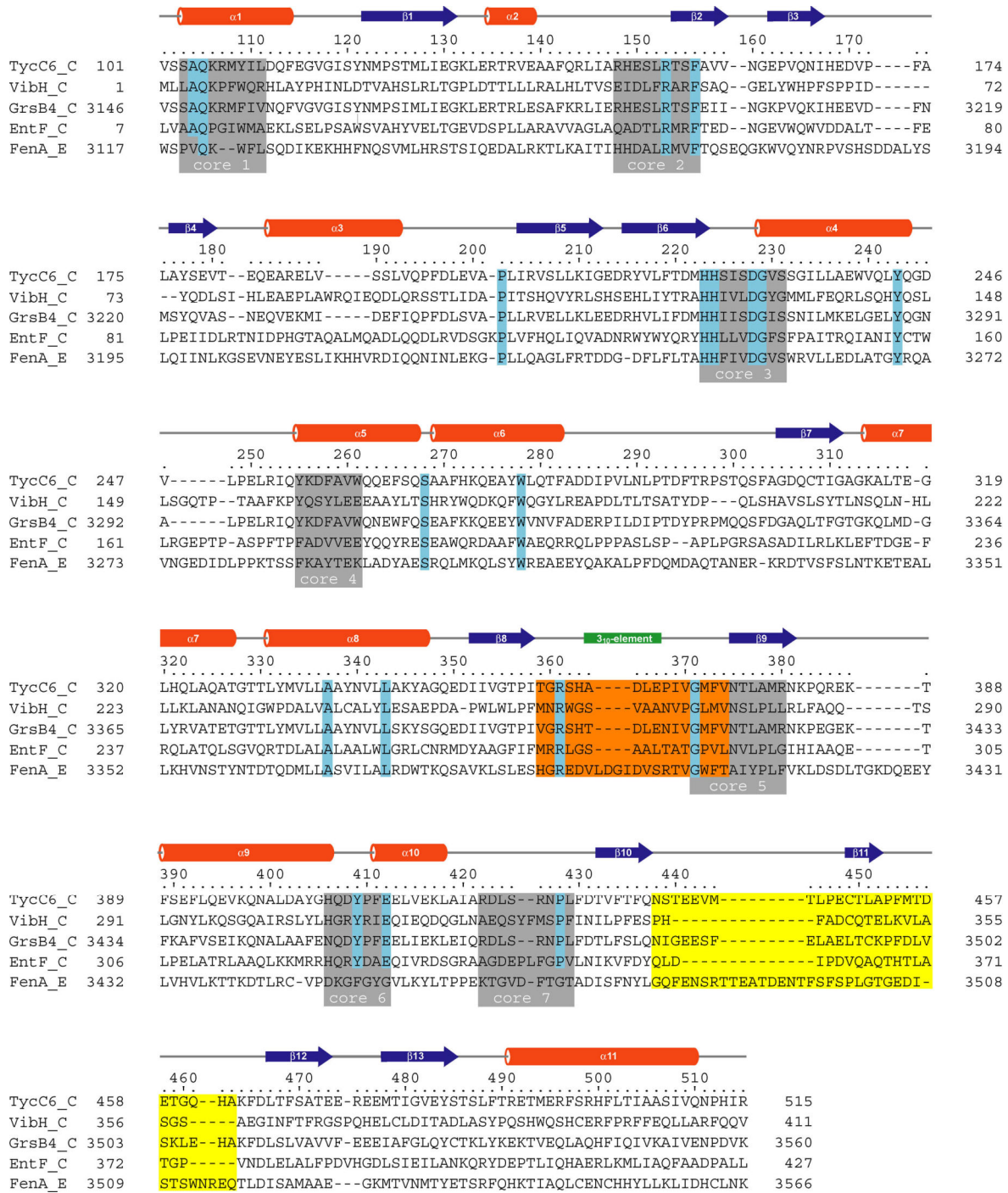
Die *bridge region* ist darüber hinaus auch derjenige Teil dieser Struktur, der die höchste Mobilität aufweist. Das zeigt sich an den erhöhten B-Faktoren. Aufgrund fehlender Elektronendichte im Bereich von T456 bis G460 konnten diese fünf Reste nur mit einer Besetzungszahl von null in der Struktur grob positioniert werden.

Anhand eines strukturbasierten Sequenzvergleichs von insgesamt 105 Kondensationsdomänen mit dem Programm 3D-COFFEE [Notredame, 2000; O'Sullivan, 2004] zeigte sich für die *bridge region* auch ein geringerer Konservierungsgrad (Abbildung 5.7). Die beobachtete erhöhte Mobilität und Sequenzvariabilität der *bridge region* erinnern stark an die *lid regions* in NRPS Thioesterase-domänen. Allerdings ließ sich auch im Falle der *bridge region* anhand der Sequenzvergleiche und der vorhandenen Strukturdaten kein Zusammenhang zur Art und chemischen Natur der Donor- oder Akzeptorsubstrate herstellen.



**Abbildung 5.7: Graphische Darstellung des Konservierungsgrades in NRPS-Kondensationsdomänen**  
 Die PCP-C Bidomäne aus zwei verschiedenen Perspektiven (links: von der Donorseite; rechts: von der Akzeptorseite). Der Konservierungsgrad eines jeden Aminosäurerestes ist farblich wiedergegeben. Blaue Reste sind stark konserviert und rote Reste sind sehr variabel. Es ist zu erkennen, dass sowohl in der Nachbarschaft des Serinrestes der PCP-Domäne als auch im aktiven Zentrum der C-Domäne ein hoher Konservierungsgrad vorliegt. Die *bridge region* ist hingegen sehr variabel.





**Abbildung 5.8: Sequence alignment von vier NRPS-Kondensationsdomänen und einer Epimerisierungsdomäne**

Vergleich der Kondensationsdomäne des Moduls 6 der Tyrocidin Synthetase TycC (*B. brevis*) mit den Kondensationsdomänen von VibH (*V. cholerae*, 19 % Identität mit TycC6), GrsB4 (*B. brevis*, 64 %) und EntF (*E. coli*, 18 %). Zum Vergleich mit den strukturell ähnlichen Epimerisierungsdomänen ist in der untersten Zeile die Epimerisierungsdomäne der Synthetase FenA (*B. subtilis*, 25 %) aufgeführt.

Die von Konz *et al.* beschriebenen, für C-Domänen charakteristischen Core-Motive sind grau unterlegt, die Sequenzen der bridge region und des floor loop sind orange bzw. gelb hervorgehoben. Weitere in C-Domänen konservierte Reste sind hellblau markiert.

In der Kristallstruktur werden neben dem Protein auch zwei Sulfationen angetroffen, die aus dem Kristallisationspuffer stammen und die wie zwei Marker vermeintlich wichtige Bindungsstellen besetzen. Das erste Sulfation liegt im Substratkanal der Kondensationsdomäne und steht über Wasserstoffbrückenbindungen in Kontakt mit den Resten G229 und H224, zwei konservierten Resten des *Core*-Motivs C3 (HHxxxDG). Eine weiterführende Betrachtung dieses Sulfations und der daraus abgeleiteten Implikationen für die Katalyse erfolgt in der Diskussion des aktiven Zentrums der Kondensationsdomäne (Kapitel 5.5, Seite 92).

Das zweite Sulfation ist an die PCP-Domäne gebunden und interagiert dort über Wasserstoffbrückenbindungen bzw. Salzbrücken mit den Haupt- und Seitenketten der Reste H42 und R45. Die Bindungsstelle befindet sich in unmittelbarer Nähe zum konservierten Serinrest S43. An diesen Rest ist in der *holo*-Form des Proteins der 4'-Phosphopantethein-Kofaktor kovalent gebunden. Bedenkt man, dass die Bindung an den Serinrest über eine Phosphorsäureesterbindung erfolgt, dann ist es möglich, dass das Sulfation in einer Tasche gebunden wird, in der üblicherweise die Phosphatgruppe des Kofaktors zum Liegen kommt.

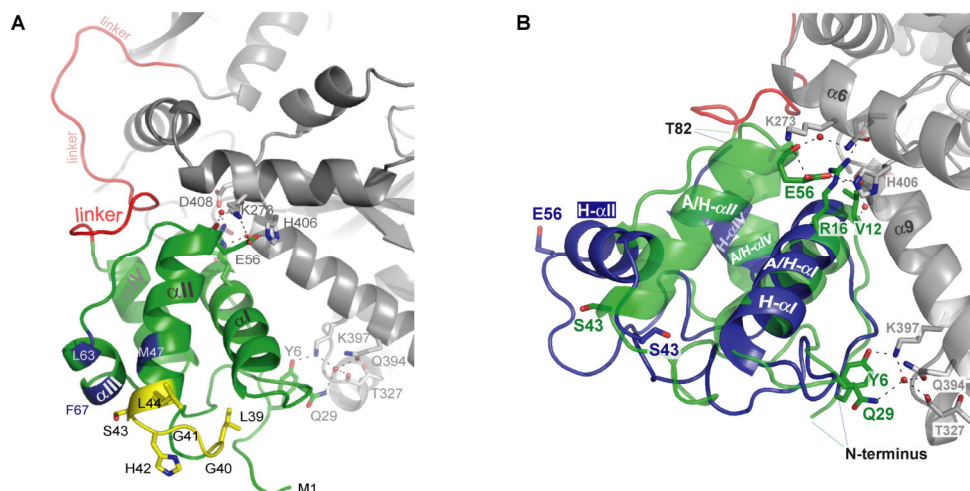


## 5.4 Wechselwirkungen zwischen der C-Domäne, ihrer donorseitigen PCP-Domäne und dem Domänen-verbindenden *Linker*-Bereich

In der Kristallstruktur zeigen beide Domänen eine relative Orientierung, bei der die katalytischen Zentren mit den Resten S43 und H224 etwa 47 Å voneinander entfernt liegen. Diese Distanz ist zu groß, als dass sie von dem 4'-Ppan-Kofaktor der PCP-Domäne überbrückt werden könnte. Daher wird davon ausgegangen, dass die in der Kristallstruktur beobachtete Anordnung nicht diejenige ist, die die beiden Domänen während der Kondensationsreaktion einnehmen.

Darüber hinaus gibt es ein weiteres Indiz dafür, dass eine andere Domänenorientierung bei der Kondensation zweier PCP-gebundener Substrate vorliegt. So hatten Lai *et al.* am Beispiel der Synthetase EntB untersucht, welche Reste der PCP-Domäne für die Interaktion mit ihrer *downstream* gelegenen Kondensationsdomäne (EntF-C) von Bedeutung sind. Die dort identifizierten Reste, deren Mutation die Interaktion mit der nachfolgenden C-Domäne störte, waren M249, F264 sowie A267. Die entsprechenden Reste der TycC5-PCP-Domäne sind M47, L63 und F67 (Abbildung 5.9 A, blau markiert), die in der Nähe des aktiven Zentrums der PCP-Domäne liegen. Allerdings ist diese Seite in der Kristallstruktur der PCP-Domäne der Kondensationsdomäne abgewandt. Die in der Kristallstruktur vorgefundenen Interaktionen zwischen den beiden Domänen verteilen sich auf zwei *Cluster*, die in Abbildung 5.9 A wiedergegeben sind.

Basierend auf der Konformation der TycC3-PCP-Domäne im H-Zustand wurde in der vorliegenden Arbeit ein Strukturmodell für die TycC5-PCP-Domäne erstellt. Dieses wurde zum Vergleich mit



**Abbildung 5.9: Wechselwirkungen zwischen PCP- und C-Domäne in der PCP-C-Kristallstruktur**

(A) Die Wechselwirkungen zwischen der PCP-Domäne (grün) und der C-Domäne (grau) konzentrieren sich auf zwei *Cluster*. Die Reste des PCP-Core-Motivs (LGGHSL) sind gelb hervorgehoben und die Reste, die nach Lai *et al.* für die Kommunikation mit der nachfolgenden C-Domäne wichtig sind, sind blau markiert. Der Linkerbereich ist rot dargestellt. (B) Überlagerung des beobachteten A/H-Zustands (grün) mit dem Modell des H-Zustands (blau). Die Konformationsänderung der PCP-Domäne bewirkt eine große Umstrukturierung, wodurch einige Wechselwirkungen aufgehoben werden.

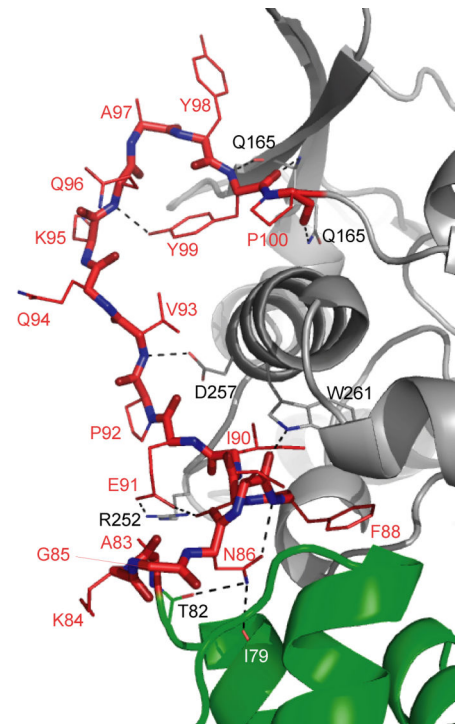
der in der Kristallstruktur beobachteten A/H-ähnlichen Struktur überlagert. Das Ergebnis ist in Abbildung 5.9B wiedergegeben und deutet darauf hin, dass die Domäne in Folge einer Konformationsänderung eine drastische Umstrukturierung erfährt. Dabei sollten auch einige in der Kristallstruktur beobachtete Wechselwirkungen zwischen der PCP- und der C-Domäne verloren gehen. Zudem zeigt die Überlagerung, dass nach wie vor über den Konformationswechsel hinaus auch eine Bewegung der PCP-Domäne als Ganzes für eine produktive Wechselwirkung mit der C-Domäne erforderlich ist.

Der 18 Aminosäurereste umfassende Linker verbindet die beiden Domänen, weist aber selbst keine Sekundärstruktur auf. Er verläuft entlang der Unterseite der C-Domäne und tritt sowohl mit der PCP-Domäne, als auch mit der C-Domäne über Wasserstoffbrückenbindungen in Wechselwirkung. Außerdem fällt auf, dass

der letzte Abschnitt des Linkers (Q94 bis P100) kaum Wechselwirkungen mit der C-Domäne eingeht. Den erhöhten B-Faktoren nach zu urteilen, ist dieser Bereich beweglich. Die Flexibilität des Linkers ist durchaus von Bedeutung, da eine statische PCP-Domäne mit ihrem etwa 20 Å langen Kofaktor nicht alle aktiven Zentren der umliegenden Domänen erreichen würde. Daher muss die gesamte PCP-Domäne ihre Position verändern. Dies erfordert einen flexiblen Linker, der die strukturelle Reorganisation der Domänen ermöglicht.

Dieser Aspekt trat auch in den Kristallstrukturen der Fettsäuresynthasen aus dem Pilz *Thermomyces lanuginosus* [Jenni, 2006] und dem Schwein [Maier, 2006; Maier, 2008] auf. Dort waren die gesamten ACP-Domänen und ihre beidseitigen Linkerbereiche in der Elektronendichte nicht definiert. Dieser Umstand ist auf die inhärente Mobilität dieser Domänen zurückzuführen, die für ihre biologische Funktion, den Transport der Intermediate zwischen den katalytischen Zentren, unerlässlich ist.

In den vergangenen Jahren wurden einige Kristallstrukturen veröffentlicht, die neue Einblicke in die räumliche Domänenanordnung und die domänenverbindenden Linkerbereiche modular aufgebauter Multidomänenenzyme gestatten. Das Terminationsmodul des Surfactin Biosynthesecusters enthält drei Interdomänen-Linker, die ähnlich zu dem hier vorliegenden Fall an den Oberflächen der Domänen liegen und teilweise in Wechselwirkungen mit den Domänen eingebunden sind [Tanovic, 2008].



**Abbildung 5.10: Detailsicht des Linkers (rot) und seiner Wechselwirkungen mit der PCP- bzw. C-Domäne**

Neue Erkenntnisse über den exakten Verlauf der Interdomänen-Linker und ihre Wechselwirkungen mit umliegenden Domänen wurden auch im Bereich der modular organisierten Typ I Fettsäure Synthasen erhalten [Maier, 2008]. In der Struktur der dimeren Fettsäure Synthase aus dem Schwein nehmen die verschiedenen Interdomänen-Linker in unterschiedlichem Ausmaß strukturgebende Funktionen wahr. Zwar verlaufen die meisten Verbindungen direkt über an der Proteinoberfläche gelegene Linker, doch bildet der 71 Aminosäuren umfassende Abschnitt zwischen der Ketosynthase- (KS) und der Malonyl-CoA-Acyl-Transferase-Domäne (MAT) eine eigenständige, strukturierte Einheit. Zwei  $\alpha$ -Helices sowie ein dreisträngiges  $\beta$ -Faltblatt wirken wie ein Puffer und unterbinden so direkte Wechselwirkungen der beiden Domänen miteinander. Der die Dehydratase- (DH) und die MAT- Domäne verbindende Linker ist 39 Aminosäurereste lang und umläuft die zentrale KS-Domäne an deren Oberfläche. Zwischen der KS- und der DH-Domäne windet sich der Linker nahe der zweizähligen Symmetrieachse einmal um sein Symmetrieäquivalent.

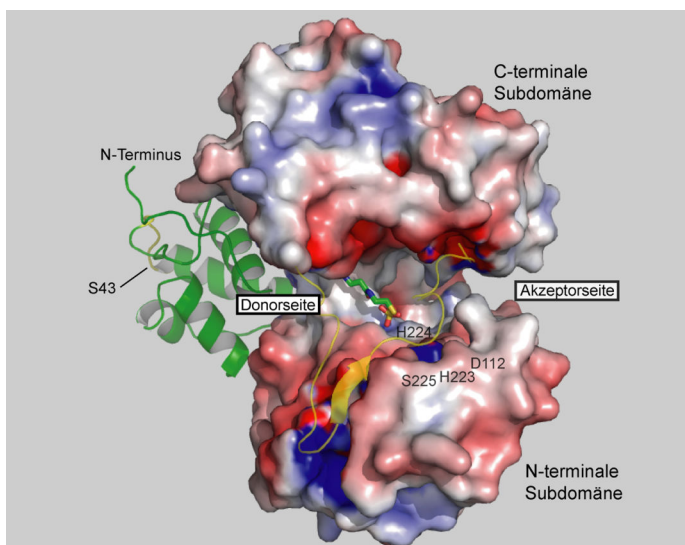
Im Gegensatz zu den zuvor genannten Beispielen der Typ I Fettsäure Synthasen sind die Interdomänen-Linker in Polyketidsynthasen auch an intermolekularen Wechselwirkungen beteiligt, die zur Dimerisierung dieser Proteine beitragen. Aus dem 6-Deoxyerythronolid B Polyketidsynthasecluster wurden die Strukturen der Ketosynthase-Acyltransferase-Bidomänen der Module 3 und 5 aufgeklärt. Beim KS-AT-Fragment des Moduls 5 lagern sich die vor den KS-Domänen liegenden,  $\alpha$ -helikalen Linker entlang der zweizähligen Symmetrieachse des Dimers aneinander [Tang, 2006]. Da der entsprechende Bereich in dem anderen kristallisierten Bidomänenprotein, dem KS-AT-Fragment des 3. Moduls, fehlt, basiert die Dimerisierung in diesem Fall allein auf den Wechselwirkungen der Ketosynthasedomänen untereinander [Tang, 2007].

## 5.5 Das aktive Zentrum der Kondensationsdomäne

Das aktive Zentrum der Kondensationsdomäne liegt in der Mitte des Substratkanals, der das Protein wie ein Graben (*canyon*) zwischen den beiden Subdomänen durchzieht und von der *bridge region* überspannt wird. Dabei wird der zweite Histidinrest des *Core*-Motivs C2 als Bezugspunkt herangezogen, da die biochemische Charakterisierung der Kondensationsdomänen der Module TycB1 [Bergendahl, 2002], EntF [Roche, 2003] und TycC6 [Schönafinger, 2003] darauf hindeutet, dass dieser Histidinrest eine entscheidende Rolle im Mechanismus der Peptidbindungsbildung spielt. Im rekombinanten Protein PCP-C entspricht dieser Rest H224. Die Seitenkette dieses Restes ragt in Richtung des Substratkanals und wäre damit in der Lage, mit den von der Donor- bzw. Akzeptorseite angereicherten Substraten zu interagieren.

In der Kristallstruktur ist im aktiven Zentrum der Kondensationsdomäne ein Sulfation gebunden. Dieses stammt aus dem Kristallisationspuffer und bildet sowohl mit der Amidfunktion des Restes G229 als auch mit der Seitenkette des katalytisch wichtigen H224 Wasserstoffbrückenbindungen aus. Darüber hinaus ist die  $\alpha$ -Helix  $\alpha 4$  so orientiert, dass ihr Dipolmoment vom aktiven Zentrum weggerichtet ist. Damit weist das positiv geladene Ende der Helix auf das im aktiven Zentrum gebundene Sulfation und wirkt stabilisierend auf dessen negative Ladung.

Eine interessante Entdeckung brachte die Überlagerung der Struktur von PCP-C mit der des Carnitin-Acetyltransferase-CoA-Komplexes (PDB: 2H3P) [Hsiao, 2004]. Das Coenzym A war so positioniert, dass es, bezogen auf die TycC6-Kondensationsdomäne, auf der Donorseite lag und in deren aktives Zentrum hineinragte. Die Sulfhydrylgruppe des 4'-Phosphopantethein-Anteils kam dabei praktisch auf dem gebundenen Sulfation zum Liegen, wie es in Abbildung 5.11 dargestellt ist.

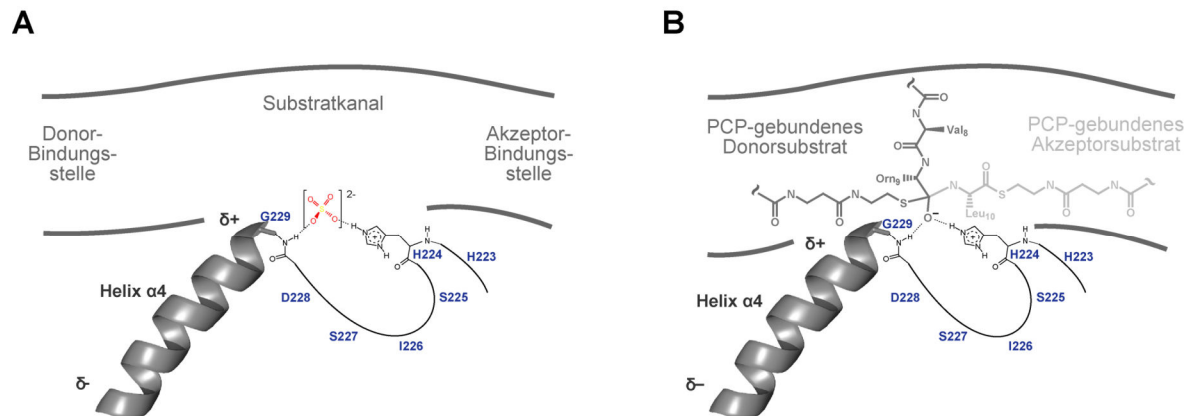


**Abbildung 5.11: Aufsicht auf die TycC6-Kondensationsdomäne mit dem 4'-Ppan-Kofaktor aus der Überlagerung mit der Struktur der Carnitin-Acetyltransferase**

Die *bridge region* ist in der Oberflächenrepräsentation ausgelassen, um den Blick auf den Substratkanal und das aktive Zentrum der Kondensationsdomäne zu gewähren. Der Verlauf der *bridge region* ist als gelbe Linie und der  $\beta$ -Strang  $\beta 11$  als gelber Pfeil dargestellt.

Die Position der 4'-Ppan-Sulphydrylgruppe stimmt fast mit der des im aktiven Zentrum gebundenen Sulfations überein.

Damit stellt sich die Frage, ob der durch das Sulfation markierten Stelle eine besondere Bedeutung zukommt. Die Wasserstoffbrückenbindungen und die elektrostatische Stabilisierung durch das Dipolmoment der  $\alpha$ -Helix  $\alpha 4$  lassen die Überlegung zu, ob das Sulfation das zwitterionische Intermediat der Kondensationsreaktion ‚imitiert‘ und im Oxyanionenloch der Kondensationsdomäne gebunden ist. Eine vergleichende Gegenüberstellung der Wechselwirkungen des Sulfations in der Kristallstruktur von PCP-C bzw. des bei einer NRPS-Kondensationsreaktion erwarteten  $sp^3$ -hybridisierten Reaktionsintermediats ist in Abbildung 5.12 skizziert.



**Abbildung 5.12: Schematische Darstellungen zur Interpretation des in der C-Domäne gebundenen Sulfations**

(A) Skizze des Substratkanals mit den in der Kristallstruktur beobachteten Wechselwirkungen des Sulfations.

(B) Skizze der erwarteten, tetraedrischen Zwischenstufe in der von TycC6-C katalysierten Kondensationsreaktion.

Möglicherweise unterstützt H224 nicht die Deprotonierung des Akzeptorsubstrats, sondern dessen nukleophilen Angriff auf das Donorsubstrat. So könnte H224 die Ausbildung des tetraedrischen Intermediats erleichtern, indem es die negative Ladung des Oxyanions stabilisiert. Diese Vorstellung einer elektrostatischen Katalyse wird durch eine mit Hilfe des H $^{++}$ -Servers [Gordon, 2005] durchgeführten, auf der Kristallstruktur der C-Domäne basierenden Berechnung der  $pK_a$ -Werte bestärkt. Diese Berechnungen ergeben für die Seitenkette von H224 einen  $pK_a$  von 11,8, was bedeutet, dass dieser Rest in seiner protonierten Form vorliegt. Dies wäre mit der elektrostatischen Katalyse vereinbar. Dann käme allerdings eine Rolle als katalytische Base nicht in Betracht.

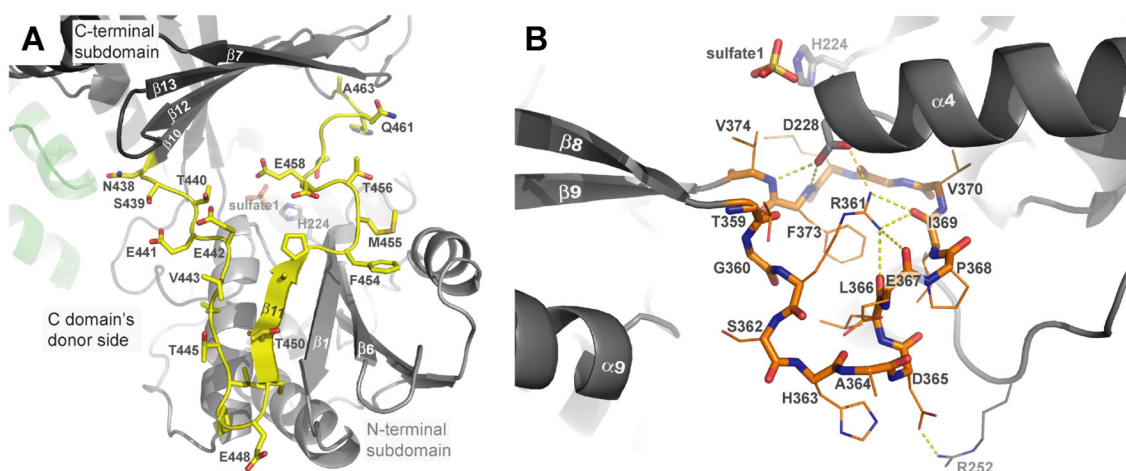
Im Übrigen stehen die Beobachtungen aus den Untersuchungen über die Rolle des Histidinrestes auch mit dem Modell der elektrostatischen Katalyse in Einklang. In den drei zuvor genannten Kondensationsdomänen TycB1-C, EntF-C und TycC6-C hatte die Mutation der entsprechenden Histidinreste jeweils die vollständige Inaktivierung der Kondensationsdomäne zur Folge. Ein anderes Verhalten zeigte sich mit der isoliert stehenden Kondensationsdomäne VibH: Nach Mutation des vergleichbaren Histidinrestes zu Alanin bzw. Valin nahm die Aktivität zwar deutlich ab, es wurde aber dennoch eine Restaktivität von etwa 10 % des Wildtyp-Proteins beobachtet [Keating, 2002]. Dem Modell der elektrostatischen Katalyse zufolge leisten die Wasserstoff-

brückenbindungen des Histidin- und Glycinrestes sowie das Dipolmoment der  $\alpha$ -Helix  $\alpha 4$  voneinander unabhängige Beiträge zur Stabilisierung der zwitterionischen Zwischenstufe. Zusammen bewirken sie eine Absenkung der Aktivierungsenergie und katalysieren so die Kondensationsreaktion. Dies schließt aber nicht aus, dass VibH einen Sonderfall darstellt, in dem in der Histidin→Alanin-Mutante die zwei verbleibenden Beiträge ausreichen und so die Kondensationsreaktion – wenn auch mit reduzierter Geschwindigkeit – auch in Abwesenheit des Histidinrestes ermöglichen. Wegen der großen Ähnlichkeit der strukturell charakterisierten Kondensationsdomänen, lässt sich dieser Unterschied auf Basis der Proteinstruktur allein nicht erklären.

### Die *bridge region* und der *floor loop*

Die Umgebung des aktiven Zentrums wird durch zwei Bereiche des Proteins bestimmt. Dabei handelt es sich um die *bridge region*, die das aktive Zentrum wie eine Brücke überspannt sowie den *floor loop*, der die Unterseite des aktiven Zentrums bildet. Diese beiden Bereiche sind Teil der C-terminalen Subdomäne und treten mit der N-terminalen Subdomäne in Wechselwirkung.

Zumindest für den *floor loop* kann davon ausgegangen werden, dass diese Wechselwirkungen für die Struktur des aktiven Zentrums und für die zur Katalyse erforderliche Positionierung der Reste (und Sekundärstrukturelemente) essentiell sind. Dafür spricht das dichte Netz von Wasserstoffbrückenbindungen und Salzbrücken zwischen den Resten des *floor loops* und den Resten



**Abbildung 5.13: Detailansicht der Umgebung des aktiven Zentrums der Kondensationsdomäne**

(A) Die *bridge region* überdacht das aktive Zentrum und interagiert mit der N-terminalen Subdomäne. (B) Der *floor loop* bildet den Boden des aktiven Zentrums. Die Wechselwirkungen konservierter Reste mit verschiedenen Bereichen der N-terminalen Subdomäne scheinen essentiell für die Aufrechterhaltung der katalytisch aktiven Konformation des aktiven Zentrums.



verschiedener angrenzender Teilbereiche der N-terminalen Subdomäne. In diesem Zusammenhang ist der konservierte Argininrest R361 zu nennen, der mit dem gleichermaßen konservierten Rest D228 des *Core*-Motivs C3 (HHxxxDG) und mit weiteren Resten des *floor loops* interagiert.

Die Mutation des entsprechenden Argininrestes in der Synthetase EntF zu Methionin (R278M) hatte zur Folge, dass die Kondensationsdomäne keine katalytische Aktivität mehr zeigte [Roche, 2003]. Da der Argininrest nicht direkt am Katalysemechanismus beteiligt ist, unterstreicht diese Beobachtung die Bedeutung des Restes für die konformationelle Integrität des aktiven Zentrums.

Die ebenfalls in der TycC6-C-Domäne beobachtete Salzbrücke zwischen D365 und R252 ist nicht in allen Kondensationsdomänen konserviert (Abbildung 5.13B). Dem *sequence alignment* (Abbildung 5.8) zufolge ist in der Kondensationsdomäne des 4. Moduls der Synthetase GrsB, GrsB4-C, eine vergleichbare Salzbrücke zwischen den Seitenketten der Reste R3297 und D3410 zu erwarten.

Die *bridge region* interagiert ebenfalls mit der N-terminalen Subdomäne. Ein kleiner Bereich bildet den  $\beta$ -Strang  $\beta$ 11, der sich an das 4-strängige  $\beta$ -Faltblatt der N-terminalen Subdomäne anlagert. Darüber hinaus kommt es zu hydrophoben Wechselwirkungen zwischen davor liegenden Resten mit Resten der  $\alpha$ -Helix  $\alpha$ 4 und dahinter liegenden Resten mit  $\beta$ -Strang  $\beta$ 6 und  $\alpha$ -Helix  $\alpha$ 3. Ein Vergleich der *bridge region* mit denen anderer Kondensationsdomänen (VibH, SrfA-C C-Domäne) bzw. strukturell verwandter Enzyme (PapA5, TRI101) zeigt eine gewisse Variabilität des Hauptkettenverlaufs und der Länge der Anlagerung an das  $\beta$ -Faltblatt. Ob die *bridge region* einen vergleichbaren Effekt auf die strukturelle Integrität des aktiven Zentrums hat oder ob sie einen Beitrag zur Substraterkennung leistet, ist momentan nicht mit Sicherheit zu sagen. Vermutlich wurden wegen dieser hohen Sequenzvariabilität diese Reste bislang in Mutationsstudien nicht untersucht. Es bleibt festzustellen, dass die *bridge region* über einen weiten Bereich (N438-G460) ausgesprochen mobil ist. Dies steht im Gegensatz zu dem *floor loop*, der wie der Kern der Kondensationsdomäne recht starr ist. Es ist daher fraglich, ob die *bridge region* somit überhaupt in der Lage ist, in vergleichbarer Weise zur Erhaltung der Konformation des aktiven Zentrums beizutragen oder nur einer Abschirmung der gebundenen Substrate bzw. Zwischenprodukte dient. In Hinblick auf ihre Sequenzvariabilität, auf die inhärente Mobilität sowie auf ihre Lage nahe dem aktiven Zentrum erinnern die *bridge regions* der Kondensationsdomänen an die *lid regions* der zuvor besprochenen Thioesterasedomänen.

### Vergleich der ribosomalen und nichtribosomalen Peptidbindungsbildung

In der nichtribosomalen Peptidsynthese wird die Ausbildung der Peptidbindungen von den Kondensationsdomänen katalysiert. Der Mechanismus der Reaktion beschreibt einen nukleophilen Angriff der  $\alpha$ -Aminogruppe des *downstream* PCP-gebundenen Aminoacylsubstrats auf den Thioester-Carboxylkohlenstoff des *upstream* PCP-gebundenen Aminoacyl- bzw. Peptidylintermediats. Daraus resultiert zuerst eine zwitterionische, tetraedrische Zwischenstufe. Mit der Lösung der C-S-Bindung zum Kofaktor der *upstream* gebundenen PCP-Domäne bildet sich die Peptidbindung aus. Während über die die Substrate betreffenden Reaktionsschritte Übereinkunft herrscht, ist es weitgehend unklar, wie die Kondensationsdomäne die Reaktion ermöglicht. Zwar wurde mittels Mutation konservierter Aminosäurereste deren Bedeutung für die Katalyse der Kondensationsreaktion nachgewiesen, wie diese Reste jedoch in das Reaktionsgeschehen eingreifen, konnte bislang nicht eindeutig geklärt werden.

Basierend auf der Tatsache, dass die Mutation des zweiten Histidinrestes des Core-Motivs C3 in verschiedenen NRPS-Kondensationsdomänen die Kondensationsaktivität aufhob, wurde lange Zeit angenommen, dass dieser Histidinrest als katalytische Base die  $\alpha$ -Aminogruppe des Akzeptorsubstrats deprotoniert und damit das Nukleophil generiert. Die beobachtete Restaktivität der H→A-Mutante der Kondensationsdomäne VibH schließt eine katalytisch essentielle Rolle des Histidinrestes aber aus.

Unter Berücksichtigung der kristallographischen Einblicke in das aktive Zentrum der TycC6-Kondensationsdomäne scheint es daher wahrscheinlicher, dass die Kondensationsdomäne die Substrate in eine für die Reaktion günstige Orientierung zueinander bringt und die in der tetraedrischen Zwischenstufe auftretenden Ladungen durch elektrostatische Wechselwirkungen stabilisiert.

In der ribosomalen Proteinsynthese erfolgt dieser Schritt im Peptidyltransferase-Zentrum der 50S Untereinheit des Ribosoms. In einem üblichen Elongationsschritt reagiert eine A-Stellen gebundene Aminoacyl-tRNA mit einer P-Stellen-gebundenen Peptidyl-tRNA. Der entscheidende Reaktionsschritt besteht in dem nukleophilen Angriff der  $\alpha$ -Aminogruppe der tRNA-gebundenen Aminosäure auf den Carboxylkohlenstoff der in der P-Stelle gebundenen Peptidyl-tRNA. Bei dieser Reaktion handelt es sich demnach um die Aminolyse der Oxoesterbindung.

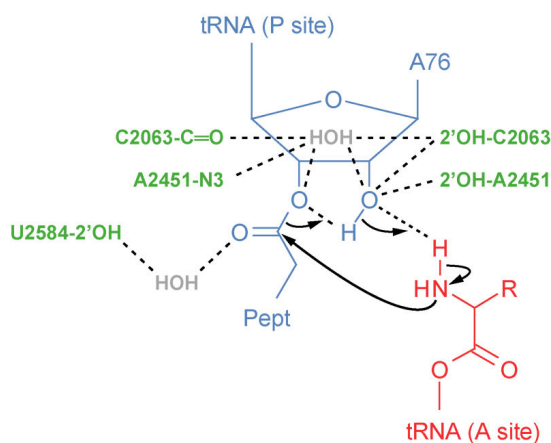
Das Ribosom und die im Ribosom stattfindenden Prozesse sind seit Jahren Gegenstand intensiver Forschung. Besonderes Interesse gilt dabei der Elongationsreaktion in der Peptidyltransferase. Die Erkenntnisse kristallographischer und biochemischer Untersuchungen haben in den vergangenen Jahren das Verständnis des zugrundeliegenden Mechanismus sowie die Rollen der an der Reaktion beteiligten Gruppen entscheidend verbessert. Mit der Lösung einer hochaufgelösten Kristallstruktur



der ribosomalen 50S Untereinheit von *Haloarcula marismortui* [Ban, 2000] wurde deutlich, dass das Peptidyltransferase-Zentrum ausschließlich aus RNA besteht. Somit sind sämtliche Ribosom-assoziierten Proteine mehr als 15 Å entfernt und greifen daher nicht direkt in den Mechanismus ein. In Anbetracht der ausgeprägten Konservierung der Nukleotidsequenz im aktiven Zentrum der Peptidyltransferase kann davon ausgegangen werden, dass diese Anordnung in allen Organismen vorgefunden wird.

Des Weiteren wurden in den Kristallstrukturen keine  $Mg^{2+}$ - oder monovalente Metallionen im aktiven Zentrum der Peptidyltransferase beobachtet [Schmeing, 2005]. Somit kann ebenfalls eine Beteiligung von Metallionen am katalytischen Geschehen ausgeschlossen werden.

Ein vorgeschlagener, aktuell diskutierter Reaktionsmechanismus der Peptidbindung bildenden Reaktion vereinigt die bislang gewonnenen Erkenntnisse. Abbildung 5.14 beschreibt die Situation im Peptidyltransferase-Zentrum unmittelbar vor dem nukleophilen Angriff der  $\alpha$ -Aminogruppe. Wie mit Hilfe der unterbrochenen Linien angedeutet, liegt ein komplexes Netz aus Wasserstoffbrückenbindungen vor. Die Reaktion verläuft vermutlich über einen sechsgliedrigen Übergangszustand und schließt die Wanderung von zwei Protonen ein.



**Abbildung 5.14: Modellvorstellung zur Reaktion im Peptidyltransferase-Zentrum nach [Beringer, 2007]**

Dieses Schema stellt die Anordnung und die Interaktionen der am Peptidyltransferase-Zentrum beteiligten Reste unmittelbar vor dem nukleophilen Angriff der  $\alpha$ -Aminogruppe dar.

Die in der P-Stelle gebundene Peptidyl-tRNA ist blau, die in der A-Stelle gebundene Aminoacyl-tRNA ist rot, Nukleotide der 23S rRNA sind grün und geordnete Wassermoleküle sind grau wiedergegeben.

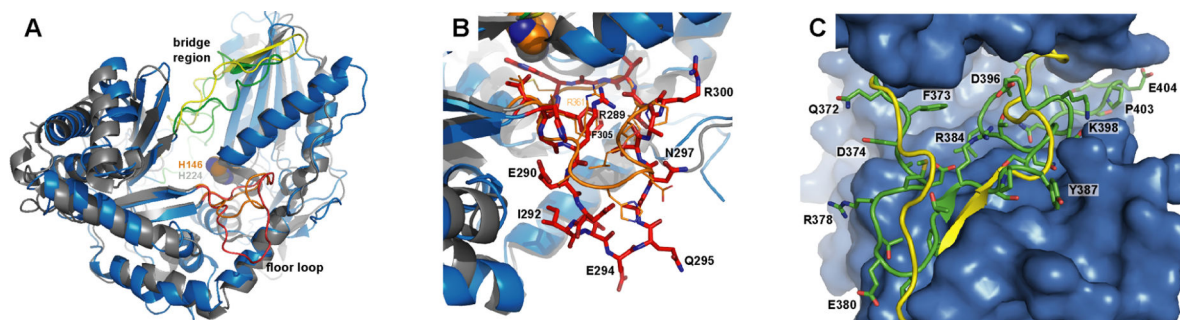
Sowohl in der ribosomalen als auch in der nichtribosomalen Peptidsynthese sind die Substrate aktiviert und liegen als Oxoester bzw. Thioester an ihre Trägermoleküle (tRNA bzw. *holo*-PCP) gebunden vor. Die Aktivierung erfolgt in einer vorgeschalteten, ATP-abhängigen Reaktion (Aminoacyl-tRNA-Synthetasen bzw. A-Domäne). In beiden Fällen schaffen die Enzyme eine Umgebung, in der der nukleophile Angriff einer  $\alpha$ -Aminogruppe erleichtert wird und dem sich die Übertragung des wachsenden Syntheseintermediats anschließt.

## 5.6 Die Struktur der TycA-Epimerisierungsdomäne

Die TycA-Epimerisierungsdomäne besteht wie die NRPS-Kondensationsdomänen aus zwei strukturell ähnlichen Subdomänen, die ebenfalls eine Verwandtschaft zum Faltungstyp der Chloramphenicol Acetyltransferase zeigen [Lewendon, 1994]. Diese strukturelle Ähnlichkeit zwischen den Kondensations- und den Epimerisierungsdomänen (und darüber hinaus auch den Zyklisierungsdomänen) war bereits anhand von Sekundärstrukturvorhersagen postuliert worden [Keating, 2002]. Allerdings reicht die Ähnlichkeit mit den C-Domänen noch weiter. Phylogenetische Studien, die sich hauptsächlich mit verschiedenen Varianten der Kondensationsdomänen beschäftigen, bestätigten die Verwandtschaft zu den Epimerisierungsdomänen [Roongsawang, 2005; Rausch, 2007].

Beide Domärentypen weisen neben ihren domänenspezifischen *Core*-Motiven ein gemeinsames konserviertes Sequenzmotiv auf, das „His-Motiv“ [C-Domäne: C3: MHHxISDG(WV)S und E-Domäne: E2: HHxISDG(WV)S]. Der zweite Histidinrest dieses Motivs (in den genannten Sequenzen unterstrichen) scheint in beiden Domänen maßgeblich an der Katalyse beteiligt zu sein, da die Mutation dieses Restes zu Alanin die Kondensations- bzw. Epimerisierungsaktivität drastisch reduziert [Stachelhaus, 1998; Stachelhaus, 2000].

In Analogie zu den Kondensationsdomänen gibt es bei der Epimerisierungsdomäne aus TycA ebenfalls eine *bridge region* und einen *floor loop*. Die beiden Strukturen sind in Abbildung 5.15A überlagert, wobei die *bridge region* und der *floor loop* farblich hervorgehoben sind. Im direkten Vergleich mit der TycC6-Kondensationsdomäne reicht der *floor loop* der Epimerisierungsdomäne weiter auf die Donorseite (Abbildung 5.15B). Ob dieser Bereich für eine spezifische Interaktion mit der PCP-Domäne oder deren 4'-Phosphopantethein-Kofaktor wichtig ist, ist bislang unklar. Dass sich jedoch die Wechselwirkungen zwischen der PCP-Domäne und der Kondensations-

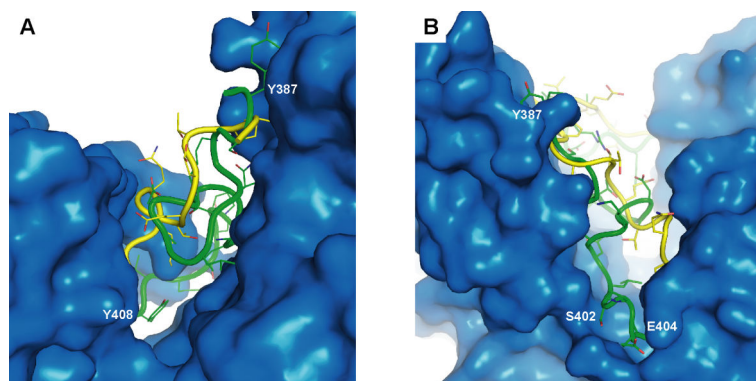


**Abbildung 5.15: Vergleich der *bridge regions* und *floor loops* in der TycC6-C- und in der TycA-E-Domäne**

(A) Totalansicht der Superposition der beiden Strukturen (C-Domäne: grau-orange-gelb; E-Domäne: blau-rot-grün). (B) Detailansicht der *floor loops*. Der Verlauf der Peptidkette variiert leicht zwischen den beiden Proteinen. (C) Detailansicht der *bridge regions*, deren Verläufe sich ebenfalls unterscheiden. Die *bridge region* der E-Domäne ist um 12 Reste länger und verschließt die ‚Rückseite‘ des aktiven Zentrums (Akzeptorseite, rechts oben).

domäne von denen der PCP-Domäne mit der Epimerisierungsdomäne unterscheiden, fiel bei der biochemischen Charakterisierung von Fusionsproteinen auf [Linne, 2001]. Infolge dieser Beobachtungen wird zwischen  $PCP^C$ - und  $PCP^E$ -Domänen unterschieden, die auch voneinander abweichende *Core*-Motive enthalten [ $PCP^C$ : GGHSI;  $PCP^E$ : GGDSI]. Die *bridge region* der Epimerisierungsdomäne umfasst 12 Reste mehr als die der TycC6-Kondensationsdomäne. Während sich der Verlauf der Hauptketten über den gesamten Bereich der *bridge region* voneinander unterscheidet, sind die stärksten Abweichungen in deren hinteren Bereich anzutreffen (T383-L401). Die Peptidkette nimmt hier einen gewundenen Verlauf und verschließt auf diese Weise die Akzeptorseite der Epimerisierungsdomäne. Die unterschiedlichen Verläufe der *bridge regions* in der C- bzw. E-Domäne sind in Abbildung 5.15C dargestellt.

Dass in E-Domänen eine längere *bridge region* vorliegt und diese ‚Insertion‘ die ‚Rückseite‘ des Substratkanals blockiert, wurde bereits aufgrund von Sequenzvergleichen und Sekundärstrukturvorhersagen postuliert [Keating, 2002]. Wie die Reste der *bridge region* der Epimerisierungsdomäne deren Akzeptorseite blockieren, veranschaulicht Abbildung 5.16 aus zwei Perspektiven.



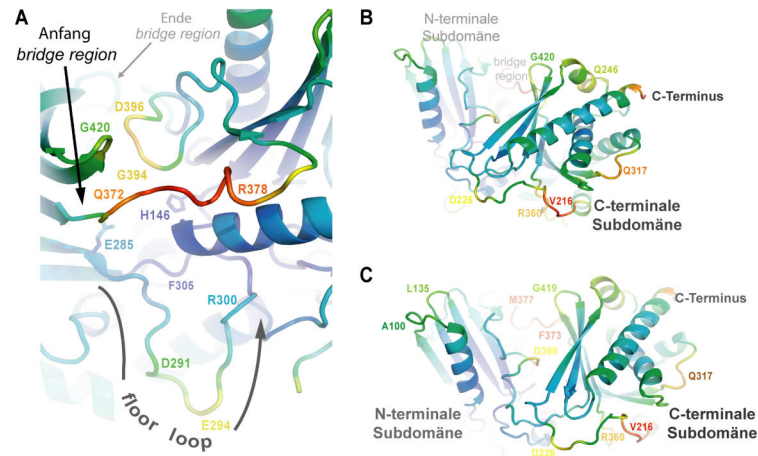
**Abbildung 5.16: Vergleich der *bridge regions* auf der Akzeptorseite des Substratkanals von TycA-E**

Darstellung der Hauptkettenverläufe der *bridge regions* von TycC6-C (gelb) und TycA-E (grün). Die blaue Proteinoberfläche stellt die TycA-E-Domäne ohne ihre *bridge region* dar. Zur Orientierung sind einige Reste von TycA-E bezeichnet. (A) Ansicht des hinteren Bereichs der *bridge region* von der Donorseite. (B) Ansicht von der Akzeptorseite.

Die *bridge region* zählt auch in der Epimerisierungsdomäne zu den beweglichsten Bereichen dieses Enzyms. Allerdings ist diese Struktur der Epimerisierungsdomäne insgesamt besser geordnet als die der PCP-C-Bidomäne, so dass die maximalen B-Faktoren der gesamten Struktur bei lediglich  $43 \text{ \AA}^2$  liegen. Dieser Wert bezieht sich auf die Hauptkettenatome der Reste Q372 bis T379, die den verhältnismäßig beweglichen Anfangsbereich der *bridge region* bilden.

Im Gegensatz dazu wird der Großteil des hinteren Bereichs der *bridge region* (T383-L401), der das aktive Zentrum auf der Akzeptorseite verschließt, durch Wechselwirkungen mit Resten des  $\beta$ -Strangs  $\beta 2$  stabilisiert. Interessanterweise liegen die zwei flexibelsten Bereiche (F373-R378 und L393-D396) direkt oberhalb des aktiven Zentrums und bilden dessen Dach. Diese sind in

Abbildung 5.17A dargestellt, wobei die Farbe den jeweiligen B-Faktor der Hauptkette wiedergibt (blau: niedriger B-Faktor [um  $10 \text{ \AA}^2$ ], rot: hoher B-Faktor [um  $40 \text{ \AA}^2$ ]). Weitere Bereiche mit erhöhten B-Faktoren liegen auf der Außen- und Unterseite der C-terminalen Subdomäne (Abbildung 5.17B und C). Eine gemäßigte Mobilität liegt auch auf der Oberseite der N-terminalen Subdomäne vor.



**Abbildung 5.17: Bewegliche Bereiche in der TycA-Epimerisierungsdomäne**

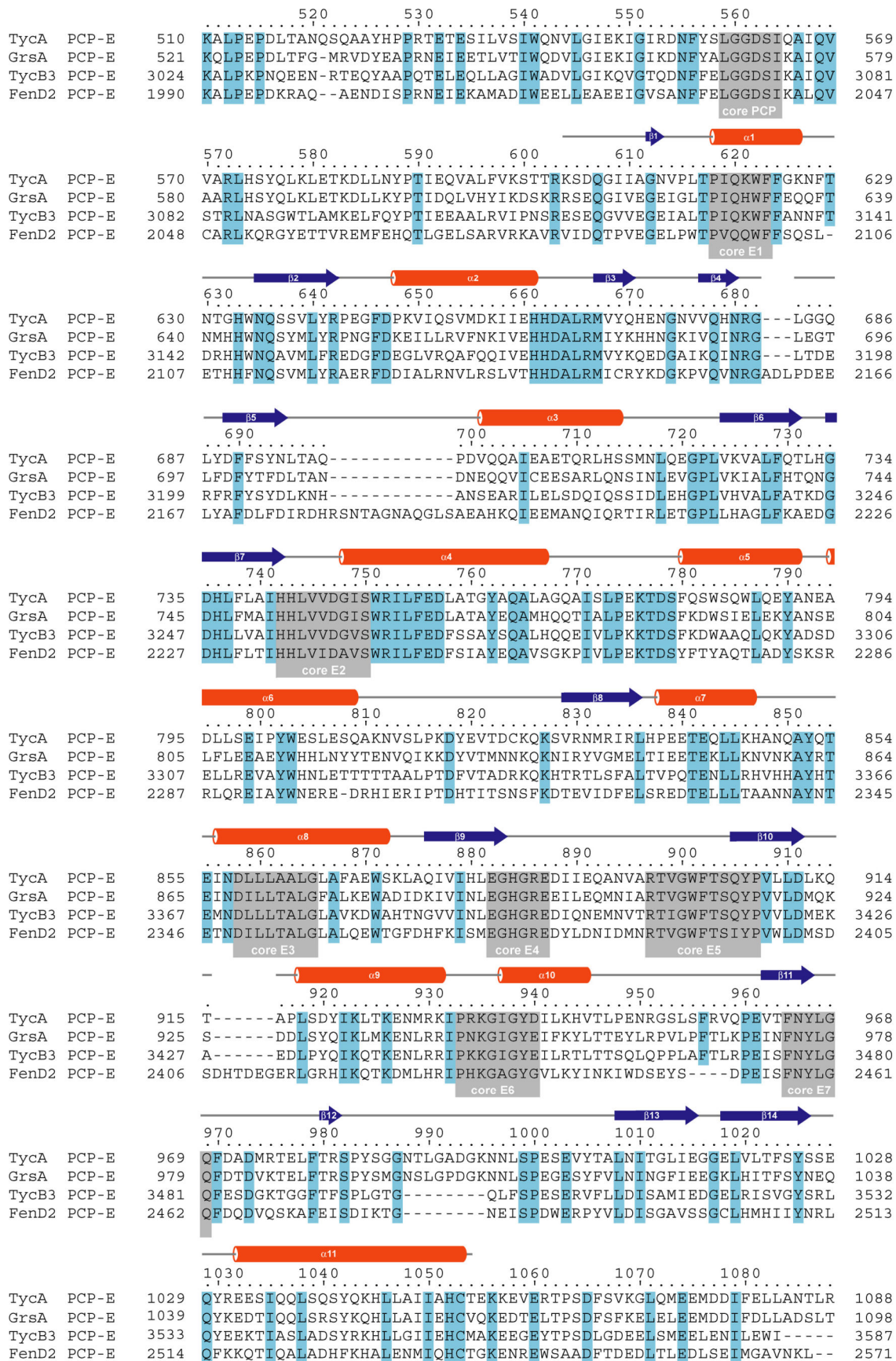
Die hier gezeigten Strukturen sind nach den B-Faktoren eingefärbt (rot kennzeichnet hohe, blau geringe Beweglichkeit). (A) Die *bridge region* liegt wie ein Dach über dem aktiven Zentrum (markiert durch die Reste H146 und E285) und weist zwei flexible Bereiche auf (Q372-E380 und G394-D396). (B) Weitere flexible Loops befinden sich auf der Außenseite der C-terminalen Subdomäne. (C) Ansicht von der Rückseite der E-Domäne. Die  $\alpha$ -Helices und  $\beta$ -Stränge verbindenden Bereiche der N-terminalen Subdomäne weisen nur mäßig erhöhte B-Faktoren auf.

Interessanterweise beschreiben Stachelhaus *et al.* beim Vergleich von internen mit C-terminalen Epimerisierungsdomänen eine geringe Sequenzähnlichkeit im Bereich der *bridge region* [Stachelhaus, 2000]. Aufgrund der Erfahrung, dass in den Kondensationsdomänen die Sequenzen in dem entsprechenden Bereich ebenfalls sehr variabel sind, erscheint es jedoch sehr spekulativ, daraus eine mögliche Beteiligung an der intermolekularen Protein-Protein-Wechselwirkung zur gegenseitigen Erkennung der Synthetas abzuleiten.

Hinsichtlich ihrer Proteinsequenzen sind die vier hier untersuchten Epimerisierungsdomänen einander recht ähnlich. Dies geht aus den Sequenzidentitäten hervor, die im direkten Vergleich jeweils 62 % (TycA:GrsA), 50 % (TycA:TycB), 38 % (TycA:FenD), 55 % (GrsA:TycB), 42 % (GrsA:FenD) bzw. 43 % (TycB:FenD) betragen.

Zum anderen wird aus dem in Abbildung 5.18 dargestellten *sequence alignment* deutlich, dass nur wenige Stellen auftreten, an denen mehrere Aminosäuren in einer oder zwei Sequenzen insertiert sind, für die es in den anderen Sequenzen keine Entsprechung gibt. Diesem *alignment* nach, besitzt die E-Domäne des Moduls TycB3 die kürzeste Sequenz.





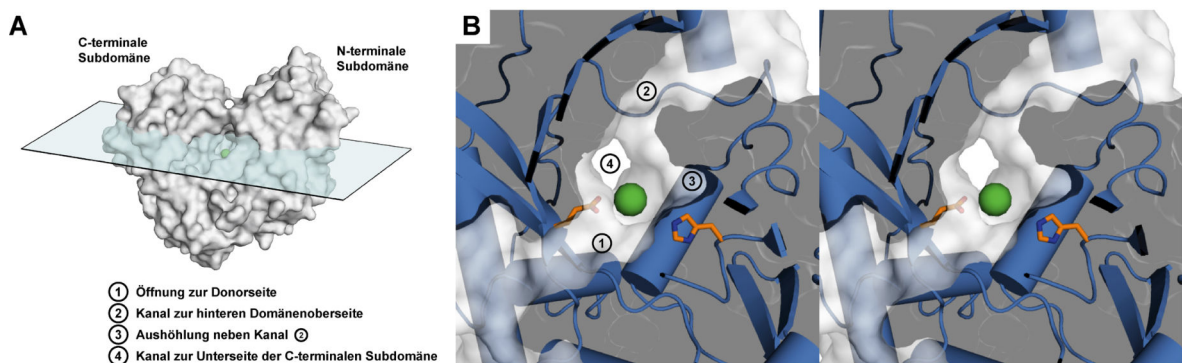
**Abbildung 5.18: Sequence alignment der in dieser Arbeit bearbeiteten NRPS-Epimerisierungsdomänen**  
 Mit dem Programm Clustal W (2.0.10) [Thompson, 1994] durchgeführtes *sequence alignment* der PCP-E-Domänen der Tyrocidin Synthetasen TycA und TycB aus *B. brevis* ATCC 8185, der Gramicidin Synthetase GrsA aus *B. brevis* ATCC 9999 sowie der Fengycin Synthetase FenD aus *B. subtilis* F29-3.

## 5.7 Das aktive Zentrum der TycA-Epimerisierungsdomäne

Das aktive Zentrum der Epimerisierungsdomäne mit dem katalytisch relevanten Histidinrest, H146, liegt mittig und verborgen zwischen den beiden Subdomänen. Der in den NRPS-Kondensationsdomänen vorgefundene, die Donor- und Akzeptorseite verbindende Kanal ist in den Epimerisierungsdomänen nur bedingt vorhanden, da er auf der Akzeptorseite durch den ausschweifend gewunden verlaufenden Loop ( $\beta$ 12- $\beta$ 13 Loop) verschlossen ist. Dies lässt den Schluss zu, dass, in Analogie zur TycC6-C Domäne, das an die *upstream* PCP-Domäne gebundene Substrat über die Donorseite zum aktiven Zentrum der Epimerisierungsdomäne gelangt. Diese Annahme ist nicht so trivial, wie sie auf den ersten Blick erscheinen mag. Schließlich muss die PCP-Domäne in einem epimerisierenden Elongationsmodul mit beiden Domänen wechselwirken, so dass die E-Domäne nicht an die Stelle der C-Domäne tritt, sondern beide Domänen nebeneinander vorliegen. Unterstützt wird diese Annahme durch die Tatsache, dass die größte, am besten zugängliche der drei vorhandenen Öffnungen des aktiven Zentrums zur Donorseite der E-Domäne führt.

Auf der diesem Zugang gegenüberliegenden Seite des aktiven Zentrums befinden sich zwei zum Teil von hydrophoben Seitenketten ausgekleidete Kanäle sowie eine Aushöhlung (Abbildung 5.19). Diese könnten die Substratbindungstaschen darstellen, wobei die apolare Umgebung im Einklang mit den Resultaten der biochemischen Charakterisierung steht. In früheren Studien hatte sich gezeigt, dass die Epimerisierungsdomäne von TycA neben Phenylalanin auch Tryptophan, Isoleucin und Valin akzeptiert, also tendenziell apolare bzw. aromatische Aminosäuren [Linne, 2001].

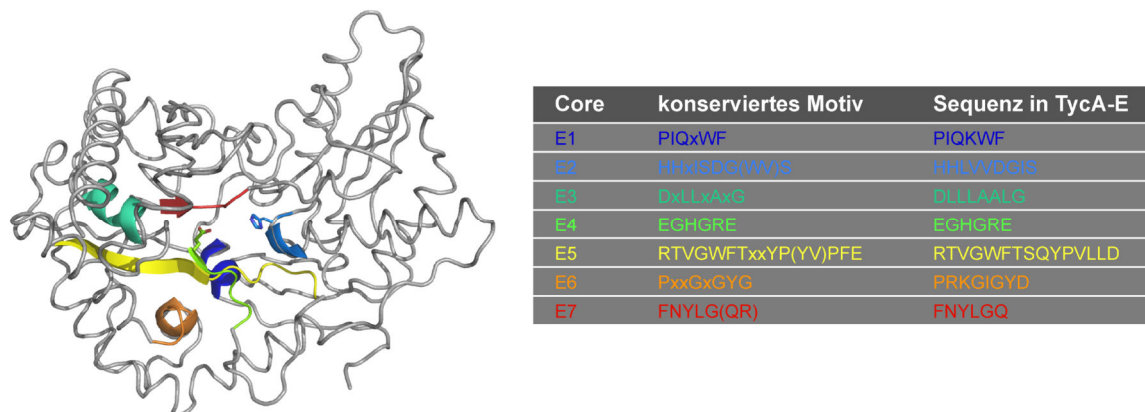
Weitere Einblicke in die Zugänge zum aktiven Zentrum der Epimerisierungsdomäne sowie eines weiteren Kanals, der vom aktiven Zentrum zur Oberseite des Proteins führt, sind im Anhang in Abbildung 8.5 gezeigt.



**Abbildung 5.19: Kanäle und Öffnungen nahe dem aktiven Zentrum der TycA-Epimerisierungsdomäne**

(A) Ansicht der Epimerisierungsdomäne von ihrer Donorseite. Die Schnitt Ebene des in (B) dargestellten Querschnitts ist angedeutet. (B) Stereoabbildung eines Querschnitts des aktiven Zentrums. Neben den Seitenketten der Reste H146 und E285 sind drei Zugänge und eine Aushöhlung nummeriert und nebenstehend beschrieben.

Darüber hinaus befinden sich in dieser auf der Akzeptorseite gelegenen Tasche einige weitere, konservierte Reste. Insgesamt verteilen sich die für Epimerisierungsdomänen charakteristischen *Core*-Motive dicht um das aktive Zentrum. Zur Veranschaulichung sind diese 7 Bereiche in Abbildung 5.20 farblich hervorgehoben.



**Abbildung 5.20: Sequenz und Position der *Core*-Motive in der Epimerisierungsdomäne aus TycA**

Lage der konservierten, für Epimerisierungsdomänen charakteristischen Sequenzmotive, die *Core*-Motive E1 bis E7 [Konz, 1999]. In der Tabelle ist neben der allgemeinen Form auch die betreffende Sequenz aus TycA wiedergegeben. Alternative Reste für eine Position sind in Klammern angegeben; „x“ beschreibt eine Position mit einer beliebigen Aminosäure.

### Strukturelle Analyse ausgewählter Mutationen in NRPS-Epimerisierungsdomänen

Eine umfassende Analyse der Epimerisierungsdomäne aus GrsA wurde von Stachelhaus *et al.* durchgeführt [Stachelhaus, 2000]. In dieser Studie wurde eine Vielzahl der konservierten Reste mit funktionellen Seitenketten durch Alanin (z. T. auch durch Serin) ersetzt und der Einfluss der Mutationen auf die Epimerisierungsreaktion untersucht. Für einen Teil der Mutation lassen sich die Auswirkungen mit Hilfe der Struktur erklären. Hier sollen nur diejenigen Reste diskutiert werden, deren Substitution in GrsA die größten Effekte auf die katalytische Aktivität zeigten. Diese Reste sind in der folgenden Tabelle zusammengefasst:

Mutation in GrsA [Stachelhaus <i>et al.</i> , 2000]	Aktivitätsbefund [Stachelhaus <i>et al.</i> , 2000]	entsprechender Rest in TycA
H753A	inaktiv	H146
D757S	inaktiv	D150
D767S	asymmetrisch beeinträchtigt	D160
E892A	beeinträchtigt	E285
R896A	beeinträchtigt	R289
Y976A	inaktiv	Y369

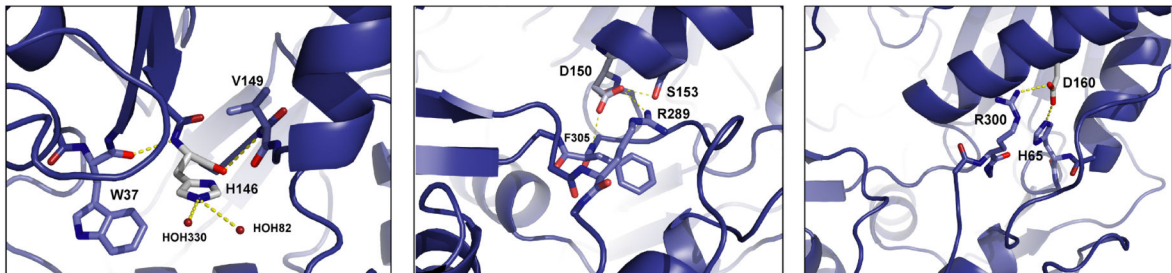
**Tabelle 6: Auswahl der von Stachelhaus *et al.* beschriebenen Mutationen in GrsA**

Die Substitution dieser Reste wirkte sich deutlich auf die Epimerisierungsaktivität aus. In der rechten Spalte sind die entsprechenden Reste der TycA-Epimerisierungsdomäne angegeben, auf die in der folgende Diskussion Bezug genommen wird.

Die Mutation des konservierten Histidinrestes **H146** reduzierte die Aktivität der Epimerisierungsdomäne um den Faktor 10 000. Der Rest liegt – ähnlich dem vergleichbaren Histidin in den Kondensationsdomänen – in der Mitte des ‚Substratkanals‘. Die Position des Histidins und die Orientierung seiner Seitenkette sind mit der Vorstellung einer Beteiligung am Katalysemechanismus sehr gut vereinbar.

Der Aspartatrest **D150** des Core-Motivs E2 ist ebenfalls von Bedeutung für die Katalyse. Die von Stachelhaus *et al.* generierte Aspartat→Alanin Mutante war unlöslich und die Aspartat→Serin-Mutante erwies sich als katalytisch inaktiv. Wie sich der Struktur entnehmen lässt, interagiert dieser Rest über seine Seitenkette mit der Hauptkette des Restes F305 sowie mit den Seitenketten der Reste S153 und R289. Bei R289 handelt es sich, analog zum mehrfach koordinierenden Rest R361 in PCP-C, um einen stark konservierten Rest des *floor loops*. Wie weiter unten bei der Behandlung von R289 beschrieben, wirkte sich dessen Substitution ebenfalls negativ auf die Aktivität der Epimerisierungsdomäne aus. Offenbar ist diese Wechselwirkung wichtig, um die zwei Subdomänen in der Umgebung des aktiven Zentrums für die Katalyse richtig zu positionieren.

Die beiden Reste, D150 und R289, sind ebenfalls in den Kondensationsdomänen konserviert und für deren Aktivität unerlässlich [Bergendahl, 2002; Keating, 2002]. Die Interaktion der beiden Reste hat offenbar auch dort Bedeutung für die strukturelle Integrität des aktiven Zentrums (vgl. Kapitel 5.5).



Umgebung und Wechselwirkungen der Aminosäurereste H146 (links), D150 (mitte) und D160 (rechts)

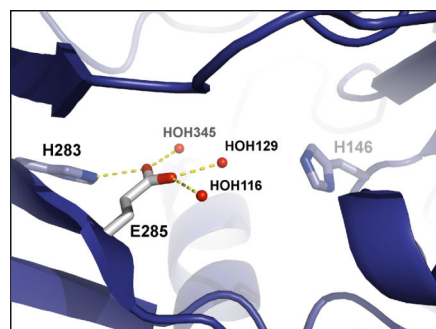
Aspartat **D160** ist Teil der Helix  $\alpha_4$ , jener Helix, die sich an das Core-Motiv E2 anschließt. Die Gruppen der Hauptkette interagieren demnach mit den Resten innerhalb der  $\alpha$ -Helix. Die Seitenkette von D160 wechselwirkt mit den Resten H65 und R300. H65 ist Bestandteil eines Sequenzmotivs (HHDALRM), welches in vielen E-Domänen vorgefunden wird, aber nicht als Core-Motiv benannt wurde [Konz, 1999]. Bei R300 handelt es sich um einen Rest des *floor loops*. R300 ist als erster Rest des Core-Motivs E5 in Epimerisierungsdomänen stark konserviert.



Warum die Substitution des Restes D160 das Enzym ausschließlich in der Richtung der D→L Reaktion beeinträchtigt, ist aus der Kristallstruktur nicht ersichtlich. Der Publikation zufolge wurde von diesem Rest nur eine Aspartat→Serin-Mutante untersucht [Stachelhaus, 2000]. Dieses rekombinante Protein mit der Domänenfolge A-PCP-E erschien in ATP-PP<sub>i</sub>-Austauschreaktionen in Hinblick auf Löslichkeit und Faltung normal.

Betrachtet man die Angaben zu den Mutanten der beiden Interaktionspartner, H65A und R300A, so weisen beide eine verminderte Löslichkeit auf und scheinen hinsichtlich ihrer Faltung beeinträchtigt. Es besteht daher die Möglichkeit, dass in der Aspartat→Serin-Mutante der Serinrest nur mit einen der beiden Reste interagieren kann und dass die daraus resultierenden, strukturellen Änderungen die Bindung oder Epimerisierung des D-Enantiomers behindern.

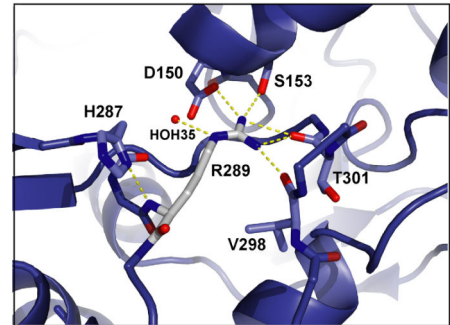
Die Mutation des Glutamats **E285** wirkte sich bei den drei von Stachelhaus *et al.* durchgeführten Analysen auf unterschiedliche Weise aus. Die Geschwindigkeit des Tritiumaustauschs zwischen dem C<sub>α</sub>-<sup>3</sup>H-markierten Substrat und dem Lösungsmittel nahm um den Faktor  $\geq 7$  ab. Auch die Geschwindigkeiten der D/L-Gleichgewichtseinstellung und der Wechselwirkung mit TycB1-ProCAT, welche zur Übertragung des D/L-Phenylalanins auf die das Modul TycB1 und anschließend zur Abspaltung von Phenylalanin-Prolin-Diketopiperazin führt, waren herabgesetzt. Dabei wirkte sich die Mutation stärker aus, wenn das Experiment von D-konfiguriertem Phenylalanin ausging. Es scheint, als sei diese Mutante des Enzyms nicht mehr in der Lage das D-Phe zu deprotonieren.



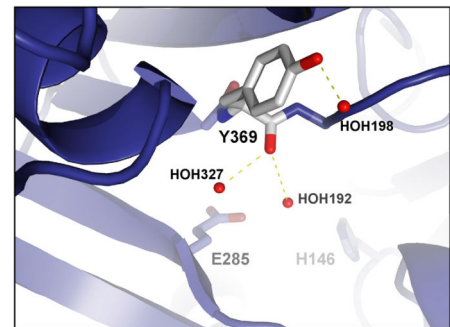
Unter der Annahme, dass ein Zwei-Basen-Mechanismus vorliegt, könnte demnach E285 eine der beiden katalytischen Basen sein. In der Kristallstruktur liegt der Rest im aktiven Zentrum und, bezogen auf H146, auf der gegenüberliegenden Seite des Substratkanals. Aufgrund dieser ‚zentralen‘ Lage ist eine Beteiligung am katalytischen Geschehen denkbar. Die Möglichkeit einer solchen Rolle wird im weiteren Verlauf der Diskussion noch behandelt. Dagegen erscheint aufgrund der Solvens-exponierten Lage der Seitenkette und der fehlenden Interaktionen eine Funktion für E285 zur Erhaltung der lokalen Proteinstruktur unwahrscheinlich.

## 5 Diskussion

Der bereits erwähnte Argininrest **R289** ist zentraler Bestandteil eines großen Netzwerks von Wasserstoffbrückenbindungen. Im Falle der TycA-E-Domäne interagiert er über seine Seitenkette mit der Hauptkette der Reste V298 und T301 und daneben mit den Seitenketten von S153 und dem oben beschriebenen D150. Die Mutation dieses Restes zu Alanin führte zu einem praktisch inaktiven Enzym. Nur in einem Experiment, dem Tritiumaustausch aus  $^3\text{H}$ -markiertem Substrat mit dem Lösungsmittel, wurde eine geringe Restaktivität festgestellt ( $\leq 1\%$  der Wildtyp-Aktivität).



Tyrosin **Y369** zeigte in dem von Stachelhaus beschriebenen Tritiumaustausch-Experiment keine messbare Aktivität. Die einzigen in der Kristallstruktur beobachteten Wechselwirkungen von Y369 sind Wasserstoffbrücken zu drei Wassermolekülen. Wegen seiner Lage an der Domänenoberfläche, neben dem Zugang zum aktiven Zentrum, ist dieser Rest eventuell für die Erkennung der PCP-Domäne oder des 4'-Ppan-Kofaktors wichtig.



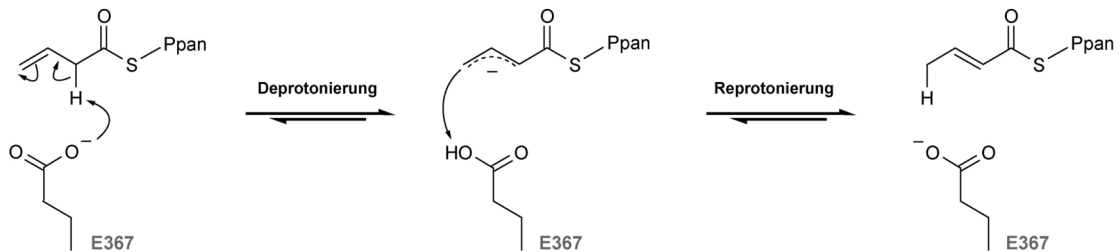
### Mechanistische Aspekte der Epimerisierungsreaktion

Über den Mechanismus der Epimerisierungsreaktion in NRPS-Epimerisierungsdomänen ist bislang wenig bekannt. In mehreren Studien wurde gezeigt, dass die Mutation des zweiten Histidinrestes des *Core*-Motivs E2 die Epimerisierungsaktivität drastisch mindert [Stachelhaus, 1998; Stachelhaus, 2000]. Man geht daher davon aus, dass dieser Histidinrest maßgeblich am Mechanismus beteiligt ist und dass er als katalytische Base fungiert. Mit den neuen, in dieser Arbeit gewonnenen Erkenntnissen über die Architektur des aktiven Zentrums und die zentrale Position des Glutamats 285, soll in diesem abschließenden Abschnitt der Arbeit über neue, mögliche Mechanismen dieser Kofaktor-unabhängigen Epimerisierung diskutiert werden.

In der Einleitung wurde bereits ausführlich dargestellt, dass eine Gruppe der bekannten Epimerasen bzw. Racemasen auf den Kofaktor Pyridoxalphosphat (PLP) angewiesen sind und dass die Bindung an diesen Kofaktor die Deprotonierung am  $C_\alpha$ -Atom begünstigt. Im Falle der Epimerisierungsdomänen aus nichtribosomalen Peptidsynthetasen ist das Aminoacyl- bzw. Peptidylsubstrat über den Kofaktor 4'-Phosphopantethein an die davorliegende PCP-Domäne gebunden. Bedingt durch die, verglichen mit Oxoestern und Amiden begrenzte mesomere Stabilisierung im Thioester, ist in diesen Substraten die Ausbildung der Enolatform erleichtert. Folglich ist die C-H-Azidität am  $C_\alpha$ -Atom erhöht und das Substrat für die folgende enzymatisch katalysierte Abstraktion des Protons aktiviert. Die Absenkung des  $pK_a$ -Wertes im aktiven Zentrum des Enzyms kann dabei mehrere Grösseneinheiten betragen. Die  $C_\alpha$ -H-Azidität einer frei in Lösung vorliegenden als Thioester gebundenen Aminosäure liegt vergleichbar mit CoA-gebundenen Fettsäuren und Ketonen bei einem  $pK_a$  von 18-20. In einer Studie an Dehydrogenasen [Vock, 1998] wurden CoA-aktivierte 3S-Acyl-Substrate eingesetzt, welche aufgrund der Abwesenheit  $\beta$ -ständiger Protonen nur Enolate bilden können. Bei diesen Untersuchungen ergab sich für die freien Acyl-CoA-Substrate eine  $C_\alpha$ -H-Azidität derselben Größenordnung ( $pK_a \approx 16.5$ ). Aus Messungen der  $C_\alpha$ -H-Azidität in den Dehydrogenase-Substrat-Komplexen wurden daraufhin  $pK_a$ -Werte von 5.2 bzw. 6.15 bestimmt. Somit erhöht allein die Bindung des Substrats an das Enzym die Azidität des  $\alpha$ -ständigen Protons um 10  $pK_a$ -Einheiten.

Eine den NRPS-Epimerisierungsdomänen ähnliche Situation, in der das Substrat als Thioester aktiviert vorliegt, findet sich in den Dehydrogenasen des Fettsäurekatabolismus. Sie katalysieren den ersten von vier repetitiven Reaktionsschritten der  $\beta$ -Oxidation zur Bildung von Acetyl-CoA aus CoA-aktivierten Fettsäuren. Das Substrat der Dehydrogenase ist dabei wie bei der Epimerisierungsdomäne als Thioester aktiviert und somit seine Azidität an der  $C_\alpha$ -Position erhöht. Es zeigte sich, dass die Dehydrogenasen neben dieser natürlichen Funktion auch die allylische Isomerisierung

zwischen  $\alpha,\beta$ - und  $\beta,\gamma$ -ungesättigten Thioestern katalysieren [Fendrich, 1982]. An dem in Abbildung 5.21 beschriebenen Beispiel der *short-chain acyl-CoA dehydrogenase* (SCAD) aus *Megasphaera elsdenii*, wird im Substrat die Position der C-C-Doppelbindung verändert. Dabei wird in der Reaktion zum Crotonylpantethein ein zu einem Thioester  $\alpha$ -ständiges Proton von einem Glutamatrest abstrahiert und dieses im zweiten Schritt an das  $\gamma$ -ständige C-Atom gebunden.



**Abbildung 5.21: Durch eine Acyl-CoA Dehydrogenase katalysierte allyliche Isomerisierung [Fendrich, 1982]**

Der Glutamatrest E367 fungiert als katalytische Base bei der reversiblen Umsetzung von Vinylacetylpanthethein (links) zu Crotonylpanthethein (rechts). Die Katalyse verläuft über ein anionisches Allylintermediat (mitte).

Interessanterweise wird auch im aktiven Zentrum der Epimerisierungsdomäne ein Glutamat, E285, angetroffen. Dieser Rest der C-terminalen Subdomäne liegt, verglichen mit dem katalytisch wichtigen Histidin H146, auf der gegenüberliegenden Seite des einseitig verschlossenen ‚Substratkanals‘ und ist etwa 7 Å von dessen Seitenkette entfernt. Eine Beteiligung dieses Restes am katalytischen Geschehen erscheint zudem plausibel, da es sich um einen invarianten Rest handelt, der Teil des *Core*-Motivs E4 ist (E4: EGHGRE; E285 unterstrichen). Wie bereits oben ausgeführt, hatte die Mutation dieses Restes in den Mutationsstudien an der Epimerisierungsdomäne von GrsA einen negativen Einfluss auf deren Epimerisierungsaktivität [Stachelhaus, 2000].

Um eine Auskunft über die Protonierungszustände bei physiologischem pH zu erhalten und damit möglicherweise eine Aussage über die Rollenverteilung im aktiven Zentrum treffen zu können, wurden die Protonierungszustände aller (de-)protonierbaren Aminosäurereste mit Hilfe des H<sup>++</sup>-Servers berechnet [Gordon, 2005]. Für den Rest H146 ergab dies einen pK<sub>a</sub>-Wert von 13.5, so dass er unter physiologischen Bedingungen, analog zum Histidin H224 der TycC6-Kondensationsdomäne, in der protonierten Form vorliegen sollte. Dies wiederum macht eine Beteiligung am Mechanismus als katalytische Base unwahrscheinlich. H146 könnte demnach entweder als katalytische Säure fungieren oder, wiederum in Analogie zur Kondensationsreaktion, das während der Reaktion auftretende Enolation stabilisieren.

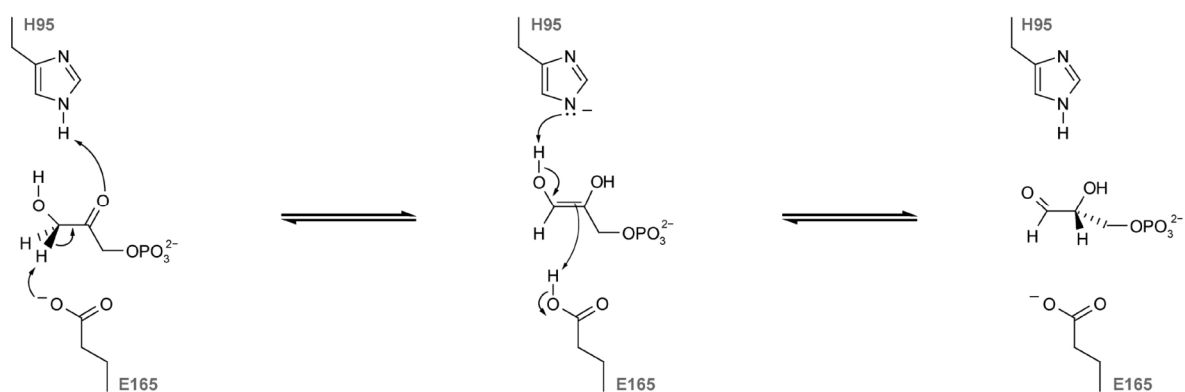
Glutamat E285 liegt den Berechnungen zufolge deprotoniert vor und käme damit als katalytische Base in Frage.

Trotz der deutlichen Unterschiede in den berechneten  $pK_a$ -Werten für Histidin H146 und Glutamat E285 (13.5 bzw. -3) ist ein Zwei-Basen-Mechanismus für die E-Domänen-katalysierte Epimerisierung durchaus möglich, da den Berechnungen ein aktives Zentrum ohne gebundenes Substrat zugrunde liegt. Die Bindung eines Substrats und die ggf. damit einhergehenden strukturellen Veränderungen könnten zu deutlichen Verschiebungen in den berechneten Aziditäten der betroffenen Gruppen führen.

Der Mechanismus der Triosephosphat-Isomerase (TIM) beschreibt die Isomerisierung als Folge zweier konzertierter Deprotonierungs-/Reprotonierungsreaktionen, an denen ein Histidin- und ein Glutamatrest als katalytische Säure bzw. Base beteiligt sind (Abbildung 5.22) [Knowles, 1991]. Es wäre durchaus möglich, dass die Reaktion im aktiven Zentrum der Epimerisierungsdomänen nach einem ähnlichen Mechanismus abläuft.

Demnach würde ein Glutamatrest als katalytische Base ein Proton abstrahieren, während ein protoniert vorliegender Histidinrest als katalytische Säure wirkte und das resultierende Enolat in die Enol-Form überführte. Auf die Situation im aktiven Zentrum der Epimerisierungsdomäne übertragen, entspräche dies einem Ein-Basen-Mechanismus mit dem Glutamatrest E285 als katalytische Base und dem protonierten Histidinrest H146 als katalytische Säure. Diese Rollenverteilung steht allerdings im Widerspruch zu dem bisher vorgeschlagenen, in der Literatur beschriebenen Modell, wonach der zweite Histidinrest des *Core*-Motivs E2 als katalytische Base fungiert und selbst an der Deprotonierung des  $C_\alpha$ -Atoms beteiligt ist.

Gegen ein Vorliegen dieses Mechanismus in NRPS-Epimerisierungsdomänen spricht allerdings, dass sich die Substitution der Glutamatreste in den beiden Enzymen unterschiedlich auf deren

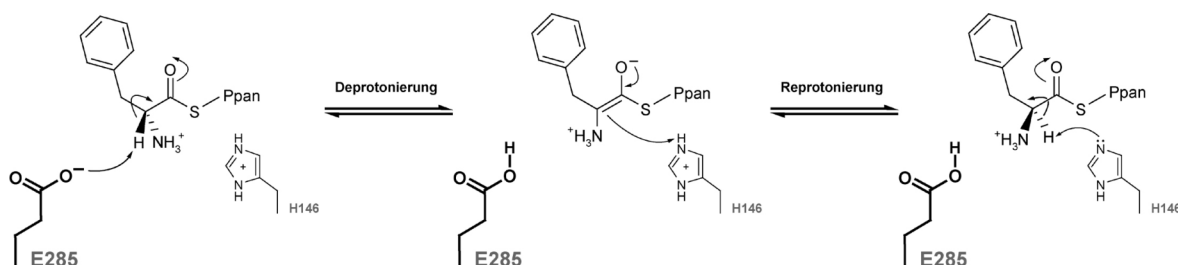


**Abbildung 5.22: Von Knowles beschriebener Reaktionsmechanismus der Triosephosphat Isomerase (TIM)**

In der Isomerisierung zwischen Dihydroxyacetonphosphat und Glyceraldehyd-3-phosphat tritt intermediär eine *cis*-Endiol-Zwischenstufe auf. Die katalytisch aktiven Aminosäurereste, Glutamat 165 und Histidin 95, agieren als katalytische Base bzw. Säure.

Aktivität auswirkt. Stachelhaus *et al.* geben für die E892A-Mutante von GrsA (E892 entspricht E285 in TycA-E) eine Restaktivität von etwa 20 % der Wildtyp-Aktivität an. Knowles nennt hingegen für analoge Mutation des Glutamatrestes in der Triosephosphat-Isomerase eine Aktivitätsminderung um den Faktor  $10^6$  [Knowles, 1991].

In der Publikation über die Analyse der Aktivität von Mutanten der GrsA-Epimerisierungsdomäne äußert Stachelhaus, dass ein Teil des beobachteten Reaktionsverhaltens mit einem Zwei-Basen-Mechanismus erklärbar wäre. Für den Fall, dass die Epimerisierungsreaktion nach einem Zwei-Basen-Mechanismus abliefe, könnten H146 und E285 als katalytisches Säure-Base-Paar agieren. Ein solches Wechselspiel ist in Abbildung 5.23 skizziert. Für den Fall, dass das Substrat so positioniert ist, dass das Glutamat die Abstraktion des  $C_\alpha$ -Protons der D-konfigurierten Aminosäure katalysiert, ließe sich auf diese Weise die von Stachelhaus *et al.* beschriebene Aktivitätsminderung der E892A-Mutante in der D→L-Richtung erklären.



**Abbildung 5.23: Modell des Zwei-Basen-Mechanismus mit H146 und E285 als katalytisches Säure-Base-Paar**

Glutamat E285 abstrahiert das  $C_\alpha$ -Proton des D-konfigurierten, Kofaktor-gebundenen Phenylalanins zum intermediären Enolat. Das Enolat wird durch das H146 zum L-Phe reprotoniert. Die Reaktion ist reversibel, so dass in der L→D-Reaktion H146 als katalytische Base und E285 als katalytische Säure auftreten.

Hinsichtlich der berichteten Inaktivität der Histidin→Alanin-Mutante H753A in GrsA [Stachelhaus, 2000] kommt für die in der vorliegenden Arbeit bearbeitete TycA-Epimerisierungsdomäne für den Histidinrest H146 folgende mögliche Funktionen in Betracht. Zum einen könnte H146, der H753 in GrsA entspricht, dafür zwingend erforderlich sein, um das Substrat richtig zu positionieren. Zum anderen könnte der Histidinrest die Enolatbildung entweder durch Protonenübertragung selbst bewirken oder diese durch elektrostatische Stabilisierung der negativen Ladung am Enolat-Sauerstoff begünstigen.

Schließlich wäre es auch denkbar, dass die Epimerisierungsdomäne lediglich die Enolatbildung katalysiert und das Enolat von einer weiteren Gruppe oder durch Wasser reprotoniert wird. Bei Vorliegen eines solchen Mechanismus würde die unterschiedliche Zugänglichkeit des Enolats für die enolatprotonierende Spezies über die Zusammensetzung des resultierenden Enantiomeren- bzw. Diastereomergemisches entscheiden.

## 5.8 Vergleich der Kondensations- und Epimerisierungsdomänen in der nichtribosomalen Peptidsynthese

Aufgrund ihrer evolutionären Verwandtschaft sind die beiden Domärentypen einander ähnlich. Betrachtet man einzig die *Core*-Motive, fällt diese Beziehung unter Umständen nicht auf. Doch spätestens die Analyse der Proteinstrukturen offenbart die ausgeprägten Gemeinsamkeiten: die konservierten Reste nahe dem aktiven Zentrum und deren konservierte Wechselwirkungen.

Interessanterweise sind die Katalysemechanismen für beide Reaktionen bislang nicht eindeutig aufgeklärt. Die bisher veröffentlichten mechanistischen Modelle beschreiben als ersten Schritt für beide Reaktionen, dass eine katalytische Base ein Proton des Substrats abstrahiert. Im Falle der Kondensationsreaktion führt dies zu einer nukleophilen  $\alpha$ -Aminogruppe, die daraufhin den Thioester des Donorsubstrats angreift. In der Epimerisierungsreaktion bildet sich eine Enol(at)-Zwischenstufe aus, die durch Reprotonierung aus dem anderen diastereotopen Halbraum in das Produkt übergeht.

Basierend auf diesen Ergebnissen wurde dem jeweils zweiten Histidinrest der *Core*-Motive C3 bzw. E2 die Rolle der katalytischen Base zugewiesen. Allerdings kommen die strukturbasierten Berechnungen der  $pK_a$ -Werte sowohl für die TycC6-Kondensations- als auch für die TycA-Epimerisierungsdomäne zu einem anderen Schluss. In beiden Fällen soll den Rechnungen zufolge der Rest protoniert vorliegen. Der protonierte Histidinrest kann die Funktion einer katalytischen Base nicht mehr erfüllen. Infolge dessen wurde für die Kondensationsdomäne das in Kapitel 5.5 beschriebene Modell der elektrostatischen Katalyse vorgeschlagen. Der Mechanismus der Epimerisierungsreaktion bleibt weiterhin ungeklärt, obwohl auch hier eine Stabilisierung der anionischen Enolatform nahe liegt. In Kapitel 5.7 wird der Mechanismus der Triosephosphat-Isomerase vorgestellt, da sie hinsichtlich ihres aktiven Zentrums und ihrer katalysierten Reaktion Ähnlichkeiten zur Epimerisierungsdomäne zeigt. Es wäre daher möglich, dass die Epimerisierung nach einem ähnlichen Mechanismus verläuft.

In Hinblick auf die Anordnung ihrer Subdomänen zueinander unterscheidet sich die in der TycA-Epimerisierungsdomäne vorgefundene Orientierung nur geringfügig von denen der strukturell bekannten Kondensationsdomänen. Dabei gibt es die größte Übereinstimmung mit der C-Domäne des Moduls TycC6 (r.m.s.d. 1.69 Å über 244  $C_\alpha$ -Atome). Zum Vergleich wurde mit VibH ein r.m.s.d.-Wert von 1.65 Å über nur 140  $C_\alpha$ -Atome und mit der C-Domäne aus SrfA-C ein r.m.s.d.-Wert von 1.51 Å über 174  $C_\alpha$ -Atome erhalten.

### **Mechanistische Ähnlichkeiten zwischen Kondensations- und Epimerisierungsreaktion**

So ähnlich wie der prinzipielle Aufbau der beiden katalytischen Domänen, so ähnlich erscheint auch der enzymatische Beitrag zu den katalysierten Reaktionen. In beiden Fällen muss das Enzym mit negativ geladenen Gruppen im Substrat interagieren. In der Stabilisierung dieser negativen Ladungen und in der Unterstützung ihrer Ausbildung liegt ein Teil der katalytischen Wirkung.

Eine Vielzahl enzymkatalysierter, PLP-unabhängiger Epimerisierungsreaktionen verlaufen über eine Enolat-Zwischenstufe. Ein solches Intermediat wird auch in der von Epimerisierungsdomänen katalysierten Reaktion erwartet, da das Substrat als Thioester aktiviert ist. Aufgrund der im Vergleich mit Oxoestern und Amiden geringeren mesomeren Stabilisierung ist die Ausbildung der Enolatform begünstigt. Es ist daher zu erwarten, dass die Epimerisierungsdomäne die Ausbildung einer Enolatform unterstützt. Dies kann beispielsweise dadurch erfolgen, dass neben der Abstraktion des jeweiligen C $_{\alpha}$ -ständigen Protons auch die resultierende negative Ladung des Enolats durch entsprechend positionierte elektrostatische Wechselwirkungen stabilisiert wird.

Im aktiven Zentrum der Kondensationsdomäne scheinen ebenfalls elektrostatische Wechselwirkungen maßgeblich an der Katalyse der Peptidbindungsbildung beteiligt zu sein. Der Katalysemechanismus verläuft über den nukleophilen Angriff einer deprotonierten  $\alpha$ -Aminogruppe auf den Thioester-Carboxylkohlenstoff des *upstream* PCP-gebundenen Aminoacyl- bzw. Peptidylintermediats. Nach Durchlaufen einer zwitterionischen, tetraedrischen Zwischenstufe führt diese Reaktion schließlich zur Spaltung der Thioesterbindung und mit der neugebildeten Peptidbindung zur Übertragung des vormals *upstream* gebundenen Syntheseintermediats.

Den Beobachtungen von Keating *et al.* an der Kondensationsdomäne VibH sowie den in dieser Arbeit gewonnen Erkenntnissen an der internen Kondensationsdomäne des Moduls TycC6 zufolge scheint die katalytische Funktion des konservierten Histidinrestes nicht in der Rolle einer katalytischen Base zu liegen. Insgesamt erscheint stattdessen eine Katalyse durch Positionierung der Substrate und durch elektrostatische Stabilisierung der intermediär auftretenden Ladungen wahrscheinlicher, wobei der Histidinrest an diesen beiden Enzym-Substrat-Wechselwirkungen beteiligt ist.

Ein entscheidender Unterschied zwischen den durch C- und E-Domänen katalysierten Reaktionen liegt in deren Kinetik. Die Studien von Stachelhaus *et al.* an der Epimerisierungsdomäne von GrsA zeigten, dass bei der frühesten Probe und selbst unter Eisbadkühlung die Epimerisierungsreaktion schon eingetreten war und sich das Gleichgewicht bereits eingestellt hatte. Die experimentell ermittelte Wechselzahl für GrsA von  $k_{\text{cat}} > 100 \text{ min}^{-1}$  kann daher nur als unterer Grenzwert angesehen werden [Stachelhaus, 2000]. Stachelhaus *et al.* untersuchten auch die von der



Kondensationsdomäne des Moduls TycB1 katalysierte Kondensationsreaktion, in der das TycA-gebundene D-Phe auf das an die PCP-Domäne des Moduls TycB1 gebundene L-Pro übertragen wird. Diese Reaktion verläuft deutlich langsamer und ergab für die TycB1-C-Domäne lediglich eine Wechselzahl von  $k_{\text{cat}} = 1.8 \text{ min}^{-1}$ .

Neben den rein enzymatischen Faktoren haben natürlich auch Unterschiede hinsichtlich der Reaktionsmechanismen und der chemischen Natur der jeweiligen Substrate einen Einfluss auf die Leichtigkeit der einzelnen Reaktionsschritte und damit auf das Voranschreiten der gesamten Kondensations- bzw. Epimerisierungsreaktion.



## 6 Ausblick

In der vorliegenden Arbeit wurden mit den Thioesterase-, Kondensations- und Epimerisierungsdomänen drei weitverbreitete Domärentypen nichtribosomaler Peptidsynthetasen strukturell untersucht. Die hier vorgestellten Erkenntnisse liefern Einblicke in die Architektur der Domänenstrukturen, in die Wechselwirkungen zwischen aufeinanderfolgenden NRPS-Domänen sowie in die Enzym-Substrat-Wechselwirkungen. Die letztgenannten Interaktionen sind erforderlich für das vollständige Verständnis der enzymatischen Katalyse. Die Kenntnis dieser Faktoren ermöglicht zum einen die Verwendung des Enzyms für Reaktionen mit nicht-natürlichen Substraten und bilden zum anderen die Grundlage des rationalen Proteindesigns, mit dessen Hilfe Enzyme auf ihre neuen Einsatzgebiete bzw. auf Substrate angepasst werden. Die Makrozyklisierung ist ein wichtiger Aspekt der nichtribosomalen Peptidsynthese, da sie erheblichen Einfluss auf die Konformation und damit auf die biologische Aktivität des Peptidprodukts hat.

Das vorgeschlagene *edge-on* Modell für den Bindungsmodus der Fengycin Thioesterase ist ein Schritt zum detaillierten Verständnis der für die Substratspezifität ursächlichen Wechselwirkungen. Dies kann jedoch nur ein Anfang sein, da nicht zuletzt, bedingt durch die ausgeprägte strukturelle Vielfalt in den NRPS-Produkten, die Erkenntnisse nur zum Teil auf Thioesterasedomänen anderer Biosynthesecuster übertragbar sein werden. Daher sind für die Abschätzung der entscheidenden Größen für Thioesterasedomänen allgemein die Lösung weiterer Thioesterase-Proteinstrukturen sowie die Bestimmung ihrer Enzym-Substrat-Wechselwirkungen auf molekularer Ebene unerlässlich.

Weitere interessante, noch offene Fragen betreffen die strukturellen Gegebenheiten in iterativ agierenden Thioesterasen und die strukturellen Unterschiede zwischen lakton- und laktambildenden Thioesterasedomänen.

Die Kondensationsdomänen zeigen ebenfalls zu einem gewissen Maß eine Substratspezifität und nehmen daher auch eine Kontrollfunktion wahr. Dies kann die Verwendung einer Kondensationsdomäne für Reaktionen mit nichtkognaten Substraten einschränken und muss bei der der Synthese neuer, nicht-natürlicher Naturstoffderivate berücksichtigt werden. Zum vollständigen Verständnis ist wiederum die Kenntnis der am Mechanismus beteiligten Gruppen erforderlich. Gegebenenfalls rückt mit dem Wissen der kritischen Faktoren auch die Möglichkeit zur gezielten Veränderung der Substratspezifität in greifbare Nähe. Momentan sind diese Details allerdings noch unbekannt. Zukünftige Studien werden hoffentlich zeigen, welche weiteren Reste in die Reaktion eingreifen, und wie die über den Kofaktor an das Enzym gebundenen Substrate in den beiden Substratbindungsstellen erkannt werden. Inwiefern wird der Kofaktor gebunden und wie viele Aminosäurereste eines *upstream* synthetisierten Peptidylsubstrats werden in der donorseitigen

Substratbindungstasche erkannt. Außerdem bleibt zu klären, ob die *bridge region* und der *floor loop* Einfluss auf die Kondensationsreaktion nehmen.

In Hinblick auf die Epimerisierung in nichtribosomalen Peptidsynthetasen ermöglichte die Strukturaufklärung der TycA-Epimerisierungsdomänen erstmals einen Einblick in die strukturelle Organisation dieser Domänen. Einigen Aminosäureresten konnte anhand der Kristallstruktur eine Funktion zugeteilt werden, für andere Reste konnte ihre direkte Beteiligung am katalytischen Geschehen ausgeschlossen werden. Während wie bei den Kondensationsdomänen auch hier die Reaktionsschritte des Substrats bekannt scheinen, sind der genaue Katalysemechanismus und die Beiträge einzelner Aminosäurereste unbekannt. So bietet die Kristallstruktur eine ideale Basis für zukünftige Mutations- und Inhibitorstudien, um den Mechanismus aufzuklären. Auch weitere kristallographische Arbeiten könnten dabei hilfreich sein. So könnten weitere Versuche die Kristallisation der verkürzten TycB3-Epimerisierungsdomäne beinhalten, um mit dieser Struktur als Suchmodell die bislang ungelöste, komplex-oligomere Kristallstruktur mittels *molecular replacement* zu lösen. Die Wechselwirkungen in diesem Oligomer könnten Aufschluss über die zur Erkennung aufeinanderfolgender Synthetasen erforderlichen, intermolekularen Wechselwirkungen geben. Es wird davon ausgegangen, dass der als  $\alpha$ -helikal vorhergesagte C-Terminus der Epimerisierungsdomäne mit einem dreisträngigen  $\beta$ -Faltblatt der Kondensationsdomäne interagiert, wie dies in der Kristallstruktur von des Terminationsmoduls SrfA-C beobachtet wurde [Tanovic, 2008]. Nach der Beobachtung, dass C- und A-Domänen eine verhältnismäßig stabile Einheit bilden, sind für Strukturstudien auch Fusionsproteine der Domänenfolge C-A-PCP\*-E ~ C-A denkbar (mit ~ als Fusionsstelle mit einem flexiblen Linker und PCP\* als PCP-Domäne, in welcher der katalytisch aktive Serinrest zu einem Alanin mutiert ist).

Daneben könnten C-terminal verkürzte PCP-E Bidomänenproteine kristallisieren und ihre Strukturen nützliche Einblicke in die Interaktionen zwischen PCP- und E-Domänen liefern. Darüberhinaus ließen sich auf diese Weise substratähnliche Inhibitoren kovalent an das Enzym binden, was vermutlich die Chance auf einen kristallographischen Nachweis erhöhen würde.

## 7 Literaturverzeichnis

- Balibar, C.J., Vaillancourt, F.H. and Walsh, C.T.** (2005) Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chem Biol*, **12**, 1189-1200.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A.** (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905-920.
- Bashford, D. and Karplus, M.** (1990) pK<sub>a</sub> of ionizable groups in proteins: Atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219-10225.
- Belshaw, P.J., Walsh, C.T. and Stachelhaus, T.** (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science*, **284**, 486-489.
- Bergendahl, V., Linne, U. and Marahiel, M.A.** (2002) Mutational analysis of the C-domain in nonribosomal peptide synthesis. *Eur J Biochem*, **269**, 620-629.
- Beringer, M. and Rodnina, M.V.** (2007) The ribosomal peptidyl transferase. *Mol Cell*, **26**, 311-321.
- Bruner, S.D., Weber, T., Kohli, R.M., Schwarzer, D., Marahiel, M.A., Walsh, C.T. and Stubbs, M.T.** (2002) Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure*, **10**, 301-310.
- Bu, X., Wu, X., Xie, G. and Guo, Z.** (2002) Synthesis of tyrocidine A and its analogues by spontaneous cyclization in aqueous solution. *Organic Letters*, **4**, 2893-2895.
- Buglino, J., Onwueme, K.C., Ferreras, J.A., Quadri, L.E. and Lima, C.D.** (2004) Crystal structure of PapA5, a phthiocerol dimycocerosyl transferase from *Mycobacterium tuberculosis*. *J Biol Chem*, **279**, 30634-30642.
- Butler, M.S. and Buss, A.D.** (2006) Natural products--the future scaffolds for novel antibiotics? *Biochem Pharmacol*, **71**, 919-929.
- CCP4** (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **50**, 760-763.
- Clugston, S.L., Sieber, S.A., Marahiel, M.A. and Walsh, C.T.** (2003) Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry*, **42**, 12095-12104.
- Conti, E., Stachelhaus, T., Marahiel, M.A. and Brick, P.** (1997) Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo J*, **16**, 4174-4183.
- Crosa, J.H. and Walsh, C.T.** (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev*, **66**, 223-249.
- Davies, D.R. and Segal, D.M.** (1971) Protein crystallization: Micro techniques involving vapor diffusion. *Methods Enzymol*, **22**, 266-269.

**De Crecy-Lagard, V., Marliere, P. and Saurin, W.** (1995) Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation. *C R Acad Sci III*, **318**, 927-936.

**DeLano, W.L.** (2002). The PyMOL molecular graphics system.

**Dieckmann, R., Lee, Y.O., van Liempt, H., von Dohren, H. and Kleinkauf, H.** (1995) Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases. *FEBS Lett*, **357**, 212-216.

**Drake, E.J., Nicolai, D.A. and Gulick, A.M.** (2006) Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. *Chem Biol*, **13**, 409-419.

**Emsley, P. and Cowtan, K.** (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.

**Eppelmann, K., Stachelhaus, T. and Marahiel, M.A.** (2002) Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry*, **41**, 9718-9726.

**Fendrich, G. and Abeles, R.H.** (1982) Mechanism of action of butyryl-CoA dehydrogenase: reactions with acetylenic, olefinic, and fluorinated substrate analogues. *Biochemistry*, **21**, 6685-6695.

**Fischbach, M.A. and Walsh, C.T.** (2006) Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev*, **106**, 3468-3496.

**Frueh, D.P., Arthanari, H., Koglin, A., Vosburg, D.A., Bennett, A.E., Walsh, C.T. and Wagner, G.** (2008) Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature*, **454**, 903-906.

**Fukuda, D.S., Du Bus, R.H., Baker, P.J., Berry, D.M. and Mynderse, J.S.** (1990) A54145, a new lipopeptide antibiotic complex: isolation and characterization. *J Antibiot (Tokyo)*, **43**, 594-600.

**Garvey, G.S., McCormick, S.P. and Rayment, I.** (2008) Structural and functional characterization of the TRI101 trichothecene 3-O-acetyltransferase from *Fusarium sporotrichioides* and *Fusarium graminearum*: kinetic insights to combating *Fusarium* head blight. *J Biol Chem*, **283**, 1660-1669.

**Gehring, A.M., Mori, I. and Walsh, C.T.** (1998) Reconstitution and characterization of the *Escherichia coli* enterobactin synthetase from EntB, EntE, and EntF. *Biochemistry*, **37**, 2648-2659.

**Gibrat, J.F., Madej, T. and Bryant, S.H.** (1996) Surprising similarities in structure comparison. *Curr Opin Struct Biol*, **6**, 377-385.

**Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S. and Onufriev, A.** (2005) H<sup>++</sup>: a server for estimating pK<sub>a</sub>s and adding missing hydrogens to macromolecules. *Nucleic Acids Res*, **33**, W368-W371.

**Hahn, M. and Stachelhaus, T.** (2004) Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proc Natl Acad Sci U S A*, **101**, 15585-15590.

- Hahn, M. and Stachelhaus, T.** (2006) Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. *Proc Natl Acad Sci U S A*, **103**, 275-280.
- Hoffmann, K., Schneider-Scherzer, E., Kleinkauf, H. and Zocher, R.** (1994) Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis. *J Biol Chem*, **269**, 12710-12714.
- Hojati, Z., Milne, C., Harvey, B., Gordon, L., Borg, M., Flett, F., Wilkinson, B., Sidebottom, P.J., Rudd, B.A., Hayes, M.A., Smith, C.P. and Micklefield, J.** (2002) Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. *Chem Biol*, **9**, 1175-1187.
- Hsiao, Y.S., Jogl, G. and Tong, L.** (2004) Structural and biochemical studies of the substrate selectivity of carnitine acetyltransferase. *J Biol Chem*, **279**, 31584-31589.
- Hubbard, B.K. and Walsh, C.T.** (2003) Vancomycin assembly: nature's way. *Angew Chem Int Ed Engl*, **42**, 730-765.
- Humphrey, W., Dalke, A. and Schulten, K.** (1996) VMD - Visual Molecular Dynamics. *J. Molec. Graph.*, **14**, 33-38.
- Jancarik, J. and Kim, S.H.** (1991) Sparse Matrix Sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.*, **24**, 409-411.
- Jenni, S., Leibundgut, M., Maier, T. and Ban, N.** (2006) Architecture of a fungal fatty acid synthase at 5 angstrom resolution. *Science*, **311**, 1263-1267.
- Jenni, S., Leibundgut, M., Boehringer, D., Frick, C., Mikolasek, B. and Ban, N.** (2007) Structure of fungal fatty acid synthase and implications for iterative substrate shuttling. *Science*, **316**, 254-261.
- Jones, D.T.** (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*, **292**, 195-202.
- Kabsch, W.** (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, **26**, 795-800.
- Keating, T.A., Marshall, C.G., Walsh, C.T. and Keating, A.E.** (2002) The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat Struct Biol*, **9**, 522-526.
- Kessler, N., Schuhmann, H., Morneweg, S., Linne, U. and Marahiel, M.A.** (2004) The linear pentadecapeptide gramicidin is assembled by four multimodular nonribosomal peptide synthetases that comprise 16 modules with 56 catalytic domains. *J Biol Chem*, **279**, 7413-7419.
- Kirkpatrick, P., Raja, A., LaBonte, J. and Lebbos, J.** (2003) Daptomycin. *Nat Rev Drug Discov*, **2**, 943-944.
- Knowles, J.R.** (1991) Enzyme catalysis: not different, just better. *Nature*, **350**, 121-124.
- Koglin, A., Mofid, M.R., Lohr, F., Schafer, B., Rogov, V.V., Blum, M.M., Mittag, T., Marahiel, M.A., Bernhard, F. and Dotsch, V.** (2006) Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science*, **312**, 273-276.

- Kohli, R.M., Trauger, J.W., Schwarzer, D., Marahiel, M.A. and Walsh, C.T.** (2001) Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry*, **40**, 7099-7108.
- Kohli, R.M., Takagi, J. and Walsh, C.T.** (2002) The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides. *Proc Natl Acad Sci U S A*, **99**, 1247-1252.
- Konz, D. and Marahiel, M.A.** (1999) How do peptide synthetases generate structural diversity? *Chem Biol*, **6**, R39-48.
- Kratzschmar, J., Krause, M. and Marahiel, M.A.** (1989) Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases. *J Bacteriol*, **171**, 5422-5429.
- Krebs, W.G. and Gerstein, M.** (2000) The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework. *Nucleic Acids Res*, **28**, 1665-1675.
- Kuo, M.-C. and Gibbons, W.A.** (1980) Nuclear Overhauser effect and cross-relaxation rate determinations of dihedral and transannular interproton distances in the decapeptide tyrocidine A. *Biophys. J.*, **32**, 807-836.
- Laemmli, U.K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C. and Walsh, C.T.** (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem Biol*, **3**, 923-936.
- Leslie, A.G.W.** (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, No. 26.
- Lewendon, A., Murray, I.A., Shaw, W.V., Gibbs, M.R. and Leslie, A.G.** (1994) Replacement of catalytic histidine-195 of chloramphenicol acetyltransferase: evidence for a general base role for glutamate. *Biochemistry*, **33**, 1944-1950.
- Lin, S., Van Lanen, S.G. and Shen, B.** (2009) A free-standing condensation enzyme catalyzing ester bond formation in C-1027 biosynthesis. *Proc Natl Acad Sci U S A*, **106**, 4183-4188.
- Linne, U. and Marahiel, M.A.** (2000) Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry*, **39**, 10439-10447.
- Linne, U., Doekel, S. and Marahiel, M.A.** (2001) Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. *Biochemistry*, **40**, 15824-15834.
- Maier, T., Jenni, S. and Ban, N.** (2006) Architecture of mammalian fatty acid synthase at 4.5 angstrom resolution. *Science*, **311**, 1258-1262.
- Maier, T., Leibundgut, M. and Ban, N.** (2008) The crystal structure of a mammalian fatty acid synthase. *Science*, **321**, 1315-1322.



- May, J.J., Kessler, N., Marahiel, M.A. and Stubbs, M.T.** (2002) Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc Natl Acad Sci U S A*, **99**, 12120-12125.
- Miethke, M. and Marahiel, M.A.** (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev*, **71**, 413-451.
- Mootz, H.D. and Marahiel, M.A.** (1997) The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J Bacteriol*, **179**, 6843-6850.
- Mootz, H.D., Schwarzer, D. and Marahiel, M.A.** (2002) Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *Chembiochem*, **3**, 490-504.
- Mülhardt, C.** (2006) Der Experimentator: Molekularbiologie/Genomics. *Elsevier Spektrum Akademischer Verlag*, **5. Auflage**.
- Mullis, K.B. and Faloona, F.A.** (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*, **155**, 335-350.
- Murray, B.E.** (2000) Problems and perils of vancomycin resistant enterococci. *Braz J Infect Dis*, **4**, 9-14.
- Murshudov, G.N., Vagin, A.A. and Dodson, E.J.** (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.
- Notredame, C., Higgins, D.G. and Heringa, J.** (2000) T-Coffee: A novel method for multiple sequence alignments. *J Mol Biol*, **302**, 205-217.
- O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G. and Notredame, C.** (2004) 3DCoffee: Combining protein sequences and structures within multiple sequence alignments. *J Mol Biol*, **340**, 385-395.
- Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. and Tucker, P.A.** (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr*, **61**, 449-457.
- Patel, H.M., Tao, J. and Walsh, C.T.** (2003) Epimerization of an L-cysteinyI to a D-cysteinyI residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from *Pseudomonas aeruginosa*. *Biochemistry*, **42**, 10514-10527.
- Perrakis, A., Morris, R. and Lamzin, V.S.** (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*, **6**, 458-463.
- Perrakis, A., Harkiolaki, M., Wilson, K.S. and Lamzin, V.S.** (2001) ARP/wARP and molecular replacement. *Acta Crystallogr D Biol Crystallogr*, **57**, 1445-1450.
- Peypoux, F., Bonmatin, J.M. and Wallach, J.** (1999) Recent trends in the biochemistry of surfactin. *Appl Microbiol Biotechnol*, **51**, 553-563.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K.** (2005) Scalable molecular dynamics with NAMD. *J. Comp. Chem.*, **26**, 1781-1802.

- Rausch, C., Hoof, I., Weber, T., Wohlleben, W. and Huson, D.H.** (2007) Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol Biol*, **7**, 78.
- Roche, E.D. and Walsh, C.T.** (2003) Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis. *Biochemistry*, **42**, 1334-1344.
- Roongsawang, N., Lim, S.P., Washio, K., Takano, K., Kanaya, S. and Morikawa, M.** (2005) Phylogenetic analysis of condensation domains in the nonribosomal peptide synthetases. *FEMS Microbiol Lett*, **252**, 143-151.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.** (1989) Molecular Cloning: A Laboratory Manual *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, NY.
- Samel, S.A.** (2004) Untersuchungen zur strukturellen Charakterisierung von Thioesterasedomänen nichtribosomaler Peptidsynthetasen. *Fachbereich Chemie*, Marburg: Philipps-Universität Marburg.
- Samel, S.A., Schoenafinger, G., Knappe, T.A., Marahiel, M.A. and Essen, L.O.** (2007) Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. *Structure*, **15**, 781-792.
- Schmeing, T.M., Huang, K.S., Strobel, S.A. and Steitz, T.A.** (2005) An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature*, **438**, 520-524.
- Schneider, T.R. and Sheldrick, G.M.** (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr*, **58**, 1772-1779.
- Schönafinger, G.** (2003) Charakterisierung von internen Kondensationsdomänen der nicht-ribosomalen Peptidsynthetase des Tyrocidins: Untersuchungen zur Akzeptanz von Substraten in cis und in trans. *Fachbereich Chemie*, Marburg: Philipps-Universität Marburg.
- Schwede, T., Kopp, J., Guex, N. and Peitsch, M.C.** (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res*, **31**, 3381-3385.
- Sheldrick, G.M.** (2002) Macromolecular phasing with SHELXE. *Zeitschrift für Kristallographie*, **217**, 644-650.
- Sieber, S.A., Walsh, C.T. and Marahiel, M.A.** (2003) Loading peptidyl-coenzyme A onto peptidyl carrier proteins: a novel approach in characterizing macrocyclization by thioesterase domains. *J Am Chem Soc*, **125**, 10862-10866.
- Sieber, S.A., Tao, J., Walsh, C.T. and Marahiel, M.A.** (2004) Peptidyl thiophenols as substrates for nonribosomal peptide cyclases. *Angew Chem Int Ed Engl*, **43**, 493-498.
- Sieber, S.A. and Marahiel, M.A.** (2005) Molecular mechanisms underlying nonribosomal peptide synthesis: Approaches to new antibiotics. *Chem Rev*, **105**, 715-738.
- Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I.A., Lesley, S.A. and Godzik, A.** (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics*, **23**, 3403-3405.
- Smith, S. and Tsai, S.C.** (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep*, **24**, 1041-1072.

- Stachelhaus, T., Mootz, H.D., Bergendahl, V. and Marahiel, M.A.** (1998) Peptide bond formation in nonribosomal peptide biosynthesis. *J Biol Chem*, **273**, 22773-22781.
- Stachelhaus, T., Mootz, H.D. and Marahiel, M.A.** (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol*, **6**, 493-505.
- Stachelhaus, T. and Walsh, C.T.** (2000) Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry*, **39**, 5775-5787.
- Stein, D.B., Linne, U. and Marahiel, M.A.** (2005) Utility of epimerization domains for the redesign of nonribosomal peptide synthetases. *Febs J*, **272**, 4506-4520.
- Stein, D.B., Linne, U., Hahn, M. and Marahiel, M.A.** (2006) Impact of Epimerization Domains on the Intermodular Transfer of Enzyme-Bound Intermediates in Nonribosomal Peptide Synthesis. *Chembiochem*, **7**, 1807-1814.
- Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R. and Morris, H.R.** (1996) The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J Biol Chem*, **271**, 15428-15435.
- Tang, G.L., Cheng, Y.Q. and Shen, B.** (2007) Chain initiation in the leinamycin-producing hybrid nonribosomal peptide/polyketide synthetase from *Streptomyces atroolivaceus* S-140. Discrete, monofunctional adenylation enzyme and peptidyl carrier protein that directly load D-alanine. *J Biol Chem*, **282**, 20273-20282.
- Tang, Y., Kim, C.-Y., Mathews, I.I., Cane, D.E. and Khosla, C.** (2006) The 2.7 Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc Natl Acad Sci U S A*, **103**, 11124-11129.
- Tanner, M.E.** (2002) Understanding nature's strategies for enzyme-catalyzed racemization and epimerization. *Acc Chem Res*, **35**, 237-246.
- Tanovic, A., Samel, S.A., Essen, L.O. and Marahiel, M.A.** (2008) Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science*, **321**, 659-663.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J.** (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 4673-4680.
- Trauger, J.W., Kohli, R.M., Mootz, H.D., Marahiel, M.A. and Walsh, C.T.** (2000) Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature*, **407**, 215-218.
- Tseng, C.C., Bruner, S.D., Kohli, R.M., Marahiel, M.A., Walsh, C.T. and Sieber, S.A.** (2002) Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase. *Biochemistry*, **41**, 13350-13359.
- Umezawa, H., Aoyagi, T., Nishikiori, T., Okuyama, A., Yamagishi, Y., Hamada, M. and Takeuchi, T.** (1986) Plipastatins: new inhibitors of phospholipase A2, produced by *Bacillus cereus* BMG302-ff67. I. Taxonomy, production, isolation and preliminary characterization. *J Antibiot (Tokyo)*, **39**, 737-744.
- Vagin, A. and Teplyakov, A.** (1997) MOLREP: an Automated Program for Molecular Replacement. *Journal of Applied Crystallography*, **30**, 1022-1025.

- Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. and Clardy, J.** (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol*, **229**, 105-124.
- Vanittanakom, N., Loeffler, W., Koch, U. and Jung, G.** (1986) Fengycin--a novel antifungal lipopeptide antibiotic produced by *Bacillus subtilis* F-29-3. *J Antibiot (Tokyo)*, **39**, 888-901.
- Vock, P., Engst, S., Eder, M. and Ghisla, S.** (1998) Substrate activation by acyl-CoA dehydrogenases: transition-state stabilization and pKs of involved functional groups. *Biochemistry*, **37**, 1848-1860.
- Volpon, L., Besson, F. and Lancelin, J.-M.** (2000) NMR structure of antibiotics plipastatins A and B from *Bacillus subtilis* inhibitors of phospholipase A<sub>2</sub>. *FEBS Lett*, **485**, 76-80.
- Wagner, B., Sieber, S.A., Baumann, M. and Marahiel, M.A.** (2006) Solvent engineering substantially enhances the chemoenzymatic production of surfactin. *Chembiochem*, **7**, 595-597.
- Walsh, C.** (2000) Molecular mechanisms that confer antibacterial drug resistance. *Nature*, **406**, 775-781.
- Wang, Y., Address, K.J., Chen, J., Geer, L.Y., He, J., He, S., Lu, S., Madej, T., Marchler-Bauer, A., Thiessen, P.A., Zhang, N. and Bryant, S.H.** (2007) MMDB: annotating protein sequences with Entrez's 3D-structure database. *Nucleic Acids Res*, **35**, D298-300.
- Weber, G., Schorgendorfer, K., Schneider-Scherzer, E. and Leitner, E.** (1994) The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr Genet*, **26**, 120-125.
- Weber, T., Baumgartner, R., Renner, C., Marahiel, M.A. and Holak, T.A.** (2000) Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure Fold. Des.*, **8**, 407-418.
- Yang, Z.R., Thomson, R., McNeil, P. and Esnouf, R.M.** (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics*, **21**, 3369-3376.
- Yeh, E., Lin, H., Clugston, S.L., Kohli, R.M. and Walsh, C.T.** (2004) Enhanced macrocyclizing activity of the thioesterase from tyrocidine synthetase in presence of nonionic detergent. *Chem Biol*, **11**, 1573-1582.

## 8 Anhang

### 8.1 Verwendete Abkürzungen

A	Adenylierungsdomäne, A-Domäne
aa	Aminosäure(n)
Amp	Ampicillin
APS	Ammoniumperoxodisulfat
ARP	Computerprogramm für die Proteinkristallographie
AT	Acyltransferase-Domäne
ATP	Adenosin-5'-triphosphat
bp	Basenpaar(e)
<i>bridge region</i>	Ein Bereich der Kondensations- und Epimerisierungsdomänen, der das aktive Zentrum „überbrückt“.
C	Kondensationsdomäne
CCP4i	Programm für die Proteinkristallographie
Chainsaw	Programm zur Erstellung von Suchmodellen für den molekularen Ersatz
CoA	Coenzym A
<i>Core-Motiv</i>	Für einen Domäentyp charakteristische, konservierte Sequenz
Cy	Zyklisierungsdomäne
DCC	Dicyclohexylcarbodiimid
ddNTP(s)	2',3'-di-Desoxyribonukleosid-5'-triphosphate der Basen A, T, G oder C
DH	Dehydratase-Domäne
dNTP	2'-Desoxyribonukleosid-5'-triphosphate der Basen A, T, G oder C
E	Epimerisierungsdomäne
EDTA	Ethylendiamintetraessigsäure
Ent	Enterobactin
ER	Enoylreductase-Domäne
ESI	<i>Electrospray Ionization</i>
EtBr	Ethidiumbromid
Fa.	Firma
F <sub>c</sub>	Strukturfaktor F berechnet anhand eines Strukturmodells
Fen	Fengycin
<i>floor loop</i>	Ein Bereich der Kondensations- und Epimerisierungsdomänen, der den Boden des aktiven Zentrums bildet.
FLP	<i>fengycin linear peptide</i> ; ein synthetisches, lineares Peptid

FLP-CoA	<i>fengycin linear peptide</i> , dessen C-Terminus über eine Thioesterbindung mit Coenzym A verknüpft ist
FLP-SNAC	<i>fengycin linear peptide</i> , dessen C-Terminus als Thioester mit N-Acetylcysteamin verbunden ist
$F_o$	Strukturfaktor F berechnet auf Basis der beobachteten Daten
$F_o - F_c$	Differenzelektronendichte
$2F_o - F_c$	Elektronendichte
FPLC	<i>Fast Protein Liquid Chromatography</i>
FT	Fourier-Transformation
HEPES	4-(2-Hydroxyethyl)-1-piperazinethansulfonsäure
HOBt	1-Hydroxytriazobenzoat
HPLC	<i>High Performance Liquid Chromatography</i>
IMAC	<i>Immobilized-Metal Affinity Chromatography</i>
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranosid
Kan	Kanamycin
kb	Kilobasenpaar(e)
KR	Ketoreduktase-Domäne
KS	Ketosynthase-Domäne
LB-Medium	Luria-Bertani Medium
<i>lid region</i>	flexibler, sequenzvariabler Bereich der Thioesterasedomänen, der z. T. das aktive Zentrum wie ein Deckel verschließen kann.
Loop	Proteinbereich ohne Sekundärstruktur
MALDI-MS	<i>matrix-assisted LASER-induced desorption ionization mass spectrometry</i>
MAT	Malonyl-CoA-Acetyltransferase-Domäne
MCS	<i>multiple cloning site</i>
MD	Moleküldynamik ( <i>molecular dynamics</i> )
MolRep	Programm zur Strukturlösung per molekularen Ersatzes ( <i>molecular replacement</i> )
mRNA	Boten-Ribonukleinsäure ( <i>messenger ribonucleic acid</i> )
NAMD	Programm zur Moleküldynamik-Simulation
Ni-NTA	Nickel(II)-Nitrilotriessigsäure
NRPS	Nichtribosomale Peptidsynthetase, Nichtribosomale Peptidsynthese
OD	Optische Dichte (ggf. mit Angabe der Wellenlänge in nm als Index)
PAGE	Polyacrylamid-Gelelektrophorese
PCP	Peptidyl Carrier Protein
PCR	Polymerase Kettenreaktion ( <i>polymerase chain reaction</i> )

PEG	Polyethylenglycol
Phaser	Programm für den molekularen Ersatz
4'-Ppan	4'-Phosphopantetheinyl
PKS	Polyketidsynthase, Polyketidsynthese
PMSF	Phenylmethylsulfonylfluorid
psig	pounds per square inch (Druckeinheit)
rpm	Umdrehungen pro Minute ( <i>revolutions per minute</i> )
RT	Raumtemperatur
SDS	Natriumdodecylsulfat ( <i>sodium dodecyl sulfate</i> )
SDS-PAGE	Natriumdodecylsulfat Polyacrylamid Gelelektrophorese
SNAC	S-N-Acetylcysteamin
Srf	Surfactin
SrfTE	Surfactin Thioesterase
SYBYL	Programm für die Bearbeitung von Molekül- und Proteinstrukturen
TBE	Tris(hydroxymethyl)aminomethan-Borsäure-Dinatrium-EDTA-Puffer
TCA	Trichloressigsäure ( <i>trichloroacetic acid</i> )
TE	Thioesterase
TEMED	N, N, N', N'-Tetraethylethylendiamin
TFA	Trifluoressigsäure ( <i>trifluoroacetic acid</i> )
TIPS	Tri- <i>iso</i> -propylsilan
TOF	<i>time of flight</i>
Tyc	Tyrocidin
v/v	Volumen pro Volumen ( <i>volume per volume</i> )
VMD	Visualisierungsprogramm der NAMD Suite
w/v	Gewicht pro Volumen ( <i>weight per volume</i> )

## 8.2 Abkürzungen proteinogener und nichtproteinogener Aminosäuren

Aminosäure	1-Buchstaben-Code	3-Buchstaben-Code
Alanin	A	Ala
Arginin	R	Arg
Asparagin	N	Asn
Asparaginsäure	D	Asp
Cystein	C	Cys
Glutamin	Q	Gln
Glutaminsäure	E	Glu
Glycin	G	Gly
Histidin	H	His
Isoleucin	I	Ile
Kynurenin	U	Kyn
Leucin	L	Leu
Lysin	K	Lys
Methionin	M	Met
Ornithin	O	Orn
Phenylalanin	F	Phe
Prolin	P	Pro
Serin	S	Ser
Threonin	T	Thr
Tryptophan	W	Trp
Tyrosin	Y	Tyr
Valin	V	Val



### 8.3 Core-Motive der katalytischen Domänen nichtribosomaler Peptidsynthetasen

Aus *sequence alignments* vieler NRPS-Gene wurden für die unterschiedlichen, katalytischen Domänen charakteristische, konservierte Sequenzabschnitte identifiziert. Diese sind wie ein Fingerabdruck des jeweiligen Domäentyps und ermöglichten es, neu sequenzierten Genen eine Domänenstruktur zuzuordnen. In Kombination mit dem von Stachelhaus entschlüsselten Code der nichtribosomalen Peptidsynthese ließen sich damit sowohl die Aminosäuresequenz, als auch möglicherweise vorhandene Modifikationen im Produktpeptid, wie beispielsweise Zyklisierung oder Epimerisierung eines Restes, vorhersagen. Die folgende Übersicht gibt die konservierten *Core-Motive* nach Konz *et al.* wieder [Konz, 1999].

Domäentyp	Core	konservierte Sequenz
<b>Adenylierungsdomäne</b>	A1	L(TS)YxEL
	A2	LKAGxAYL(VL)P(LI)D
	A3	LAYxxYTSG(ST)TGxPKG
	A4	FDxS
	A5	NxYGPTE
	A6	GELxlxGxG(VL)ARGYL
	A7	Y(RK)TGDL
	A8	GRxDxQVKIRGxRIELGEIE
	A9	LPxYM(IV)P
	A10	NGK(VL)DR
<b>PCP-Domäne</b>	PCP	DxFFxxLGG(HD)S(LI)
<b>Kondensationsdomäne</b>	C1	SxAQxR(LM)(WY)xL
	C2	RHExLRTxF
	C3	MHHxISDG(WV)S
	C4	YxD(FY)AVW
	C5	(IV)GxFVNT(QL)(~)xR
	C6	(HN)QD(YV)PFE
	C7	RDxSRNPL
<b>Zyklisierungsdomäne</b>	Z1	FPL(TS)xxQxAYxxGR
	Z2	RHx(IM)L(PAL)x(ND)GxQ
	Z3	(DNR)4xDxxS
	Z3	(LI)Pxx(PAL)x(LPF)P
	Z4	(TS)(PA)3x(LAF)6x(IVT)LxxW
	Z5	(GA)(DQN)FT
	Z6	P(IV)VF(TA)SxL
	Z7	QV(x(LI)Dx(QH)11xW(DYF)

Fortsetzung auf der nächsten Seite

Domänentyp	Core	konservierte Sequenz
<b>Epimerisierungsdomäne</b>	E1	PIQxWF
	E2	HHxISDG(WV)S
	E3	DxLLxAxG
	E4	EGHGRE
	E5	RTVGWFTxxYP(YV)PFE
	E6	PxxGxGYG
	E7	FNYLG(QR)
<b>Thioesterasedomäne</b>	TE	G(HY)SxG
<b>N-Methylierungsdomäne</b>	M1	VL(DE)GxGxG
	M2	NELSxYRYxAV
	M3	VEsSxARQxGxLD
<b>Reduktasedomäne</b>	R1	V(LF)(LV)TG(AV)(TN)G(YF)LG
	R2	V3xVRA
	R3	GDL
	R4	VYPYxxLRx(PL)NVxxT
	R5	GYxxSKWxxE
	R6	RPG
	R7	LExx(VI)GFLxxP

Alternative Aminosäurereste für eine Position sind in Klammern zusammengefaßt. „x“ symbolisiert eine beliebige Aminosäure und Zahlen geben die Anzahl der beliebigen Aminosäurereste wieder („3x“ entspricht „xxx“)

## 8.4 Verwendete Kristallisationslösungen

Die folgenden Tabellen geben Auskunft über die Zusammensetzungen der Kristallisationslösungen. Es wurden zum einen die Kits *Classics Suite*, *Classics Lite Suite*, *Cryo Suite*, *MBC I Suite*, *MBC II Suite* und *Anions Suite* der Firma NEXTAL verwendet. Diese wurden später von der Firma QIAgen vertrieben, auf deren Homepage auch die Zusammensetzungen dieser jeweils 96 Lösungen umfassenden Screens angegeben sind ([www.qiagen.com](http://www.qiagen.com)).

Daneben wurden ebenfalls die Kits *Basic Kit for Protein Crystallization*, *Extension Kit for Protein Crystallization*, *Cryo Kit for Protein Crystallization*, *Basic Kit for Membrane Proteins* der Firma Sigma eingesetzt. Sie bestehen jeweils aus 50 Lösungen und ihre Zusammensetzung, die auf den Ergebnissen von Jancarik und Kim [Jancarik, 1991] basieren, sind untenstehend aufgeführt.

**Sigma Basic Kit for Protein Crystallization**

Lösung	Salz	Puffer	Präzipitans
1	Ca-chloride 0.02M	Na-acetate (pH 4.6) 0.1M	2-Methyl-2,4-pentanediol 30%
2	K-,Na-tartrate 0.4M		
3	NH <sub>4</sub> -dihydrogenphosphate 0.4M		
4	NH <sub>4</sub> -sulfate 2.0M	TRIS-HCl (pH 8.5) 0.1M	
5	Na-citrate 0.2M	HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentanediol 30%
6	Mg-chloride 0.2M	TRIS-HCl (pH 8.5) 0.1M	PEG 4000 30%
7	Na-acetate 1.4M	Na-cacodylate (pH 6.5) 0.1M	
8	Na-citrate 0.2M	Na-cacodylate (pH 6.5) 0.1M	2-Propanol 30%
9	NH <sub>4</sub> -acetate 0.2M	Na-citrate (pH 5.6) 0.1M	PEG 4000 30%
10	NH <sub>4</sub> -acetate 0.2M	Na-acetate (pH 4.6) 0.1M	PEG 4000 30%
11	NH <sub>4</sub> -dihydrogenphosphate 1.0M	Na-citrate (pH 5.6) 0.1M	
12	Mg-chloride 0.2M	HEPES Na-salt (pH 7.5) 0.1M	2-Propanol 30%
13	Na-citrate 0.2M	TRIS-HCl (pH 8.5) 0.1M	PEG 400 30%
14	Ca-chloride 0.2M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 28%
15	NH <sub>4</sub> -sulfate 0.2M	Na-cacodylate (pH 6.5) 0.1M	PEG 8000 30%
16	Li-sulfate 1.5M	HEPES Na-salt (pH 7.5) 0.1M	
17	Li-sulfate 0.2M	TRIS-HCl (pH 8.5) 0.1M	PEG 4000 30%
18	Mg-acetate 0.2M	Na-cacodylate pH (6.5) 0.1M	PEG 8000 20%
19	NH <sub>4</sub> -acetate 0.2M	TRIS-HCl (pH 8.5) 0.1M	2-Propanol 30%
20	NH <sub>4</sub> -sulfate 0.2M	Na-acetate (pH 4.6) 0.1M	PEG 4000 25%
21	Mg-acetate 0.2M	Na-cacodylate (pH 6.5) 0.1M	2-Methyl-2,4-pentanediol 30%
22	Na-acetate 0.2M	TRIS-HCl (pH 8.5) 0.1M	PEG 4000 30%
23	Mg-chloride 0.2M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 30%
24	Ca-chloride 0.2M	Na-acetate (pH 4.6) 0.1M	2-Propanol 20%
25	Na-acetate 1M	Imidazole (pH 6.5) 0.1M	
26	NH <sub>4</sub> -acetate 0.2M	Na-citrate (pH 5.6) 0.1M	2-Methyl-2,4-pentanediol 30%
27	Na-citrate 0.2M	HEPES Na-salt (pH 7.5) 0.1M	2-Propanol 20%
28	Na-acetate 0.2M	Na-cacodylate pH (6.5) 0.1M	PEG 8000 30%
29	K-,Na-tartrate 0.8M	HEPES Na-salt (pH 7.5) 0.1M	
30	NH <sub>4</sub> -sulfate 0.2M		PEG 8000 30%
31	NH <sub>4</sub> -sulfate 0.2M		PEG 4000 30%
32	NH <sub>4</sub> -sulfate 2M		
33	Na-formiate 4M		
34	Na-formiate 2M	Na-acetate (pH 4.6) 0.1M	
35	K-/Na-dihydrogenphosphate 0.8M	HEPES Na-salt (pH 7.5) 0.1M	
36		TRIS-HCl (pH 8.5) 0.1M	PEG 8000 8%
37		Na-acetate (pH 4.6) 0.1M	PEG 4000 8%
38	Na-citrate 1.4M	HEPES Na-salt (pH 7.5) 0.1M	
39	NH <sub>4</sub> -sulfate 2.0M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 2%
40		Na-citrate (pH 5.6) 0.1M	2-Propanol 20%, PEG 4000 20%
41		HEPES Na-salt (pH 7.5) 0.1M	2-Propanol 10%, PEG 4000 20%
42	K-dihydrogenphosphate 0.05M		PEG 8000 20%
43			PEG 1500 30%
44	Mg-formiate 0.2M		
45	Zn-acetate 0.2M	Na-cacodylate (pH 6.5) 0.1M	PEG 8000 18%
46	Ca-acetate 0.2M	Na-cacodylate (pH 6.5) 0.1M	PEG 8000 18%
47	NH <sub>4</sub> -sulfate 2.0M	Na-acetate (pH 4.6) 0.1M	
48	NH <sub>4</sub> -dihydrogenphosphate 2.0M	TRIS-HCl (pH 8.5) 0.1M	
49	Li-sulfate 1.0M		PEG 8000 2%
50	Li-sulfate 0.5M		PEG 8000 15%

**Sigma Extension Kit for Protein Crystallization**

Lösung	Salz	Puffer	Präzipitans
1	Na-chloride 2M	PEG 6000 10%	
2	Na-chloride 0.5M	Mg-chloride 0.01M	CTAB 0.01M
3			Ethylene glycol 25%
4			Dioxane 35%
5	NH <sub>4</sub> -sulfate 2M	2-Propanol 5%	
6		Imidazole (pH 7.0) 1M	
7			PEG 1000 10%, PEG 8000 10%
8	Na-chloride 1.5M		Ethanol 10%
9	Na-chloride 2M	Na-acetate (pH 4.6) 0.1M	
10	Na-chloride 0.2M	Na-acetate (pH 4.6) 0.1M	2-Methyl-2,4-pentanediol 30%
11	Co-chloride 0.01M	Na-acetate (pH 4.6) 0.1M	1,6-Hexanediol 1M
12	Cd-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	PEG 400 30%
13	NH <sub>4</sub> -sulfate 0.2M	Na-acetate (pH 4.6) 0.1M	PEG MME 2000 30%
14	K-,Na-tartrate 0.2M	Na-citrate (pH 5.6) 0.1M	NH <sub>4</sub> -sulfate 2M
15	Li-sulfate 1M	Na-citrate (pH 5.6) 0.1M	NH <sub>4</sub> -sulfate 0.5M
16	Na-chloride 0.5M	Na-citrate (pH 5.6) 0.1M	Polyethylenimin 2%
17		Na-citrate (pH 5.6) 0.1M	tert.-Butanol 35%
18	Fe(III)-chloride 0.01M	Na-citrate (pH 5.6) 0.1M	Jeffamine M-600 10%
19		Na-citrate (pH 5.6) 0.1M	1,6-Hexanediol 2.5M
20		MES (pH 6.5) 0.1M	Mg-sulfate 1.6M
21	Ka-,Na- dihydrogenphosphate, 0.1M each	MES (pH 6.5) 0.1M	Na-chloride 2M
22		MES (pH 6.5) 0.1M	PEG 20'000 12%
23	NH <sub>4</sub> -sulfate 1.6M	MES (pH 6.5) 0.1M	Dioxane 10%
24	Cs-chloride 0.05M	MES (pH 6.5) 0.1M	Jeffamine M-600 30%
25	Co-chloride 0.01M	NH <sub>4</sub> -sulfate 1.8M	MES (pH 6.5) 0.1M
26	NH <sub>4</sub> -sulfate 0.2M	MES (pH 6.5) 0.1M	PEG MME 5000 30%
27	Zn-sulfate 0.01M	MES (pH 6.5) 0.1M	PEG MME 550 25%
28		Na-citrate (pH 6.5) 1.6M	
29	NH <sub>4</sub> -sulfate 0.5M	HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentanediol 30%
30		HEPES Na-salt (pH 7.5) 0.1M	MPD 5%, PEG 6000 10%
31		HEPES Na-salt (pH 7.5) 0.1M	Jeffamine M-600 20%
32	Na-chloride 0.1M	HEPES Na-salt (pH 7.5) 0.1M	NH <sub>4</sub> -sulfate 1.6M
33		HEPES Na-salt (pH 7.5) 0.1M	NH <sub>4</sub> -formate 2.0M
34	Cd-sulfate 0.05M, Na-acetate 1.0M	HEPES Na-salt (pH 7.5) 0.1M	
35		HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentanediol 70%
36		HEPES Na-salt (pH 7.5) 0.1M	Na-chloride 4.3M
37		HEPES Na-salt (pH 7.5) 0.1M	Ethylene glycol 8%, PEG 8000 10%
38		HEPES Na-salt (pH 7.5) 0.1M	PEG 10'000 20%
39	Mg-chloride 0.2M	TRIS-HCl (pH 8.5) 0.1M	1,6-Hexanediol 3.4M
40		TRIS-HCl (pH 8.5) 0.1M	tert.-Butanol 25%
41	Li-sulfate 1M, Ni(II)-chloride 0.01M	TRIS-HCl (pH 8.5) 0.1M	
42	NH <sub>4</sub> -sulfate 1.5M	TRIS-HCl (pH 8.5) 0.1M	Glycerol 12%
43	NH <sub>4</sub> -phosphate 0.2M	TRIS-HCl (pH 8.5) 0.1M	2-Methyl-2,4-pentanediol 50%
44		TRIS-HCl (pH 8.5) 0.1M	Ethanol 20%
45	Ni(II)-chloride 0.01M	TRIS-HCl (pH 8.5) 0.1M	PEG MME 2000 20%
46	Na-chloride 0.1M	Bicine (pH 9.0) 0.1M	PEG MME 550 20%
47	Mg-chloride 2M	Bicine (pH 7.5) 0.1M	
48		Bicine (pH 9.0) 0.1M	Dioxane 2%, PEG 20'000 10%
49	Mg-chloride 0.1 M	TRIS-HCl (pH 8.5) 0.1M	PEG 20000 15%
50			PEG 20000 20 %

**Sigma Cryo Kit for Protein Crystallization**

Lösung	Salz	Puffer	Präzipitans
1	Ca-chloride 0.02M	Na-acetate (pH 4.3) 0.1M	2-Methyl-2,4-pentandiol 30%
2	K-,Na-tartrate 0.26M		Glycerol 35%
3	NH <sub>4</sub> -phosphate 0.26M		Glycerol 35%
4	NH <sub>4</sub> -sulfate 1.5M	TRIS-HCl (pH 8.5) 0.075M	Glycerol 25%
5	Na-citrate 0.2M	HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentandiol 30%
6	Mg-chloride 0.16M	TRIS-HCl (pH 8.5) 0.08M	PEG 4000 24%, Glycerol 20%
7	Na-acetate 0.98M	Na-cacodylate (pH 6.5) 0.07M	Glycerol 30%
8	Na-citrate 0.14M	Na-cacodylate (pH 6.5) 0.07M	2-Propanol 21%, Glycerol 30%
9	NH <sub>4</sub> -acetate 0.17M	Na-citrate (pH 5.6) 0.085M	PEG 4000 25.5%, Glycerol 15%
10	NH <sub>4</sub> -acetate 0.17M	Na-acetate (pH 4.6) 0.085M	PEG 4000 25.5%, Glycerol 15%
11	NH <sub>4</sub> -dihydrogenphosphate 0.7M	Na-citrate (pH 5.6) 0.07M	Glycerol 30%
12	Mg-chloride 0.18M	HEPES Na-salt (pH 7.5) 0.09M	2-Propanol 27%, Glycerol 10%
13	Na-citrate 0.2M	TRIS-HCl (pH 8.5) 0.1M	PEG 400 30%
14	Ca-chloride 0.19M	HEPES Na-salt (pH 7.5) 0.095M	PEG 400 26.6%, Glycerol 5%
15	NH <sub>4</sub> -sulfate 0.17M	Na-cacodylate (pH 6.5) 0.085M	PEG 8000 25.5%, Glycerol 15%
16	Li-sulfate 1.125M	HEPES Na-salt (pH 7.5) 0.075M	Glycerol 25%
17	Li-sulfate 0.17M	TRIS-HCl (pH 8.5) 0.085M	PEG 4000 25.5%, Glycerol 15%
18	Mg-acetate 0.16M	Na-cacodylate (pH 6.5) 0.08M	PEG 8000 16%, Glycerol 20%
19	NH <sub>4</sub> -acetate 0.16M	TRIS-HCl (pH 8.5) 0.08M	2-Propanol 24%, Glycerol 20%
20	NH <sub>4</sub> -sulfate 0.16M	Na-acetate (pH 4.6) 0.08M	PEG 4000 20%, Glycerol 20%
21	Mg-acetate 0.2M	Na-cacodylate (pH 6.5) 0.1M	2-Methyl-2,4-pentandiol 30%
22	Na-acetate 0.17M	TRIS-HCl (pH 8.5) 0.085M	PEG 4000 25.5%, Glycerol 15%
23	Mg-chloride 0.2M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 30%
24	Ca-chloride 0.14M	Na-acetate (pH 4.6) 0.07M	2-Propanol 14%, Glycerol 30%
25	Na-acetate 0.7M	Imidazole (pH 6.5) 0.07M	Glycerol 30%
26	NH <sub>4</sub> -acetate 0.2M	Na-citrate (pH 5.6) 0.1M	2-Methyl-2,4-pentandiol 30%
27	Na-citrate 0.14M	HEPES Na-salt (pH 7.5) 0.07M	2-Propanol 14%, Glycerol 30%
28	Na-acetate 0.17M	Na-cacodylate (pH 6.5) 0.085M	PEG 8000 25.5%, Glycerol 15%
29	K-,Na-tartrate 0.52M	HEPES Na-salt (pH 7.5) 0.065M	Glycerol 35%
30	NH <sub>4</sub> -sulfate 0.17M	PEG 8000 25.5%	Glycerol 15%
31	NH <sub>4</sub> -sulfate 0.17M	PEG 4000 25.5%	Glycerol 15%
32			NH <sub>4</sub> -sulfate 1.5M, Glycerol 25%
33			Na-formate 3.6M, Glycerol 10%
34	Na-formate 1.4M	Na-acetate (pH 4.6) 0.07M	Glycerol 30%
35	Na-,K-dihydrogenphosphate 0.6M each	HEPES Na-salt (pH 7.5) 0.075M	Glycerol 25%
36		TRIS-HCl (pH 8.5) 0.065M	PEG 8000 5.2%, Glycerol 35%
37		Na-acetate (pH 4.6) 0.07M	PEG 4000 5.6%, Glycerol 30%
38	Na-citrate 1.26M	HEPES Na-salt (pH 7.5) 0.09M	Glycerol 10%
39	NH <sub>4</sub> -sulfate 1.7M	HEPES Na-salt (pH 7.5) 0.085M	PEG 400 1.7%, Glycerol 15%
40		Na-citrate (pH 5.6) 0.095M	2-Propanol 19%, PEG 4000 19%, Glycerol 5%
41		HEPES Na-salt (pH 7.5) 0.085M	2-Propanol 8.5%, PEG 4000 17%, Glycerol 15%
42	K-dihydrogenphosphate 0.04M		PEG 8000 16%, Glycerol 20%
43			PEG 1500 24%, Glycerin 20%
44	Mg-formate 0.1M		Glycerol 50%
45	Zn-acetate 0.16M	Na-cacodylate (pH 6.5) 0.08M	PEG 8000 14.4%, Glycerol 20%
46	Ca-acetate 0.16M	Na-cacodylate (pH 6.5) 0.08M	PEG 8000 14.4%, Glycerol 20%
47	NH <sub>4</sub> -sulfate 1.6M	Na-acetate (pH 4.6) 0.08M	Glycerol 20%
48	NH <sub>4</sub> -dihydrogenphosphate 1.6M	TRIS-HCl (pH 8.5) 0.08M	Glycerol 20%
49	Li-sulfate 0.8M		PEG 8000 1.6%, Glycerol 20%
50	Li-sulfate 0.4M		PEG 8000 12%, Glycerol 20%

**Sigma Basic Kit for Membrane Proteins**

Lösung	Salz	Puffer	Präzipitans
1	Na-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	2-Methyl-2,4-pentanediol 12%
2	Zn-acetate 0.1M	Na-acetate (pH 4.6) 0.1M	PEG 4000 12%
3	NH <sub>4</sub> -sulfate 0.2M	Na-acetate (pH 4.6) 0.1M	PEG 4000 10%
4	Na-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	2-Propanol 12%
5		Na-acetate (pH 4.6) 0.1M	PEG 4000 12%
6		Na-acetate (pH 4.6) 0.1M	NH <sub>4</sub> -sulfate 1M
7		Na-acetate (pH 4.6) 0.1M	Mg-sulfate 1M
8	Mg-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	PEG 400 18%
9	Li-sulfate 0.1M	Na-acetate (pH 4.6) 0.1M	NH <sub>4</sub> -dihydrogenphosphate 1M
10	Na-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	PEG 6000 12%
11	Mg-chloride 0.1M	Na-acetate (pH 4.6) 0.1M	PEG 6000 12%
12	Na-chloride 0.1M	Na-citrate (pH 5.6) 0.1M	PEG 400 18%
13	Li-sulfate 0.1M	Na-citrate (pH 5.6) 0.1M	PEG 4000 12%
14		Na-citrate (pH 5.6) 0.1M	2-Propanol 10%
15	Na-chloride 0.1M	Na-citrate (pH 5.6) 0.1M	2-Methyl-2,4-pentanediol 12%
16		Na-citrate (pH 5.6) 0.1M	Mg-sulfate 1M
17	Na-chloride 0.1M	Na-citrate (pH 5.6) 0.1M	PEG 4000 12%
18	Li-sulfate 0.1M	Na-citrate (pH 5.6) 0.1M	PEG 6000 12%
19	Mg-chloride 0.1M	Na-citrate (pH 5.6) 0.1M	2-Methyl-2,4-pentanediol 4%
20	Na-chloride 0.1M	Na-citrate (pH 5.6) 0.1M	
21	Li-sulfate 0.1M	Na-citrate (pH 5.6) 0.1M	PEG 400 4%
22	NH <sub>4</sub> -sulfate 1M	ADA (pH 6.5) 0.1M	
23	Li-sulfate 0.1M	ADA (pH 6.5) 0.1M	2-Propanol 2%, PEG 4000 12%
24	NH <sub>4</sub> -dihydrogenphosphate 1M	ADA (pH 6.5) 0.1M	
25	Mg-chloride 0.1M	ADA (pH 6.5) 0.1M	PEG 6000 12%
26		ADA (pH 6.5) 0.1M	2-Methyl-2,4-pentanediol 12%
27	Li-sulfate 0.1M	ADA (pH 6.5) 0.1M	Mg-sulfate 1M
28	Li-sulfate 0.3M	ADA (pH 6.5) 0.1M	PEG 400 4%
29	KNaHPO <sub>4</sub> 1M , NH <sub>4</sub> -sulfate 0.1M	HEPES Na-salt (pH 7.5) 0.1M	
30	Na-chloride 0.1M	HEPES Na-salt (pH 7.5) 0.1M	PEG 4000 10%
31	Mg-chloride 0.1M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 18%
32	K-,Na-tartrate 1M	HEPES Na-salt (pH 7.5) 0.1M	
33	NH <sub>4</sub> -sulfate 0.1M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 18%
34	NH <sub>4</sub> -sulfate 0.1M	HEPES Na-salt (pH 7.5) 0.1M	PEG 4000 10%
35	Na-citrate 0.1M	HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentanediol 12%
36	Na-citrate 1M	HEPES Na-salt (pH 7.5) 0.1M	
37	Mg-sulfate 0.6M	HEPES Na-salt (pH 7.5) 0.1M	PEG 400 4%
38	Mg-sulfate 0.6M	HEPES Na-salt (pH 7.5) 0.1M	2-Methyl-2,4-pentanediol 4%
39	K-,Na-tartrate 0.1M, Li-sulfate 0.1M	HEPES Na-salt (pH 7.5) 0.1M	
40	Na-chloride 0.15M	TRIS-HCl (pH 8.0) 0.1M	PEG 6000 12 %
41	Na-chloride 0.1M	TRIS-HCl (pH 8.0) 0.1M	2-Methyl-2,4-pentanediol 12%
42	Li-sulfate 0.1M	TRIS-HCl (pH 8.5) 0.1M	2-Methyl-2,4-pentanediol 12%
43	KNaHPO <sub>4</sub> 1M, (NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> 0.1M	TRIS-HCl (pH 8.5) 0.1M	
44	Na-acetate 0.1M	TRIS-HCl (pH 8.5) 0.1M	
45	Na-chloride 0.1M	TRIS-HCl (pH 8.5) 0.1M	
46	(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> 0.1M	TRIS-HCl (pH 8.5) 0.1M	PEG 6000 12%
47	K-,Na-tartrate 0.1M, Mg-sulfate 0.4M	TRIS-HCl (pH 8.5) 0.1M	
48	Li-sulfate 0.2M	TRIS-HCl (pH 8.5) 0.1M	
49	NH <sub>4</sub> -sulfate 0.5M	TRIS-HCl (pH 8.5) 0.1M	
50	Na-citrate 0.1M	TRIS-HCl (pH 8.5) 0.1M	PEG 400 5%

## 8.5 Kristallisation von NRPS-Epimerisierungsdomänen

In diesem Abschnitt werden Erfahrungen und Hypothesen zur Kristallisation der Epimerisierungsdomänen diskutiert. Schließlich werden die Gründe dargelegt, die zur Entfernung der COM-Domänen und damit zu einer neuen, lösbaren Kristallform führten.

Zunächst konnten ausschließlich mit der isolierten Epimerisierungsdomäne der Tyrocidin Synthetase B (TycB) Kristalle passabler Größe erhalten werden. Da später auch von der analogen PCP-E-Bidomäne (TycB3-PCP-E) kleinere, verwachsene Kristalle beobachtet wurden, war davon ausgegangen worden, dass diese Epimerisierungsdomäne – verglichen mit den anderen getesteten – eine bessere Neigung zur Kristallisation zeigt. Ein Sequenzvergleich der vier untersuchten Epimerisierungsdomänen und eine Suche nach Sequenzhomologen deuteten darauf hin, dass die Epimerisierungsdomäne aus TycB die kürzeste NRPS-E-Domäne ist. Die Sequenzen der anderen Domänen enthalten an einer oder mehreren Stellen kurze Insertionen. Die Kristallisation von NRPS-Thioesterasedomänen hatte zuvor die Erfahrung mit sich gebracht, dass ausgedehnte, mobile Loop-Regionen die Kristallisation beeinträchtigen. In Anlehnung daran schien es als sei die Epimerisierungsdomäne des Moduls TycB3 kristallisierbar, da ihre kompakte Sequenz für die Kristallisation am besten geeignet war.

### 8.5.1 Kristallisation C-terminal verkürzter NRPS-Epimerisierungsdomänen

Die vier ursprünglich untersuchten Epimerisierungsdomänen stellen jeweils den C-Terminus ihrer Synthetase dar. Damit sind sie voraussichtlich maßgeblich an den intermolekularen Wechselwirkungen beteiligt, die die gegenseitige Erkennung aufeinanderfolgender Synthetasen des Clusters bewirken. Anhand von Mutationsstudien wurde nachgewiesen, dass die endständigen Sequenzbereiche beider Synthetasen für die gegenseitige Erkennung essentiell sind [Hahn, 2004; 2006]. Nach diesem Modell wechselwirken diese zwei  $\alpha$ -helikalen Bereiche, die als *communication mediating domains* (COM Domänen) bezeichnet wurden.

Spätere strukturelle Daten [Tanovic, 2008] führten zu dem Vorschlag eines anderen Modells, das als Helix-Hand-Modell bezeichnet wird. Demzufolge interagiert der  $\alpha$ -helikale COM-Bereich am C-Terminus der einen, *upstream* gelegenen Synthetase mit einem dreisträngigen  $\beta$ -Faltblatt an der Oberfläche der anderen, *downstream* gelegenen Synthetase.<sup>2</sup> Entgegen dem Modell von Hahn *et al.*

---

<sup>2</sup> Interessanterweise ist ein vergleichbares  $\beta$ -Faltblatt auch in der C-Domäne des Moduls TycC6 enthalten. Da es sich hierbei um ein internes Modul handelt, kann dieses  $\beta$ -Faltblatt entweder noch vorhanden sein, obwohl es nicht erforderlich wäre oder eine andere, synthetasen-interne Wechselwirkung eingehen, um beispielsweise die Domänen richtig zueinander zu positionieren.

sind hierbei aber auch spätere Reste an der Ausbildung des die Wechselwirkungen vermittelnden  $\beta$ -Faltblatts beteiligt.

Die in der Kristallstruktur von SrfA-C angetroffene Interaktion wird mit dem natürlichen Protein nicht beobachtet. Stattdessen war in dem zur Kristallisation verwendeten, rekombinanten Protein an dessen C-Terminus eine vektor-kodierte Helix enthalten (*Myc* Epitop). Diese ähnelt hinsichtlich ihrer Aminosäuresequenz und ihrer Amphiphilie der C-terminalen COM-Helix der vorangehenden Synthetase SrfA-B.

Es war aufgefallen, dass alle bearbeiteten, rekombinanten Epimerisierungsdomänen während der Aufreinigung als Monomere vorlagen (bestimmt anhand des Elutionsvolumens mit einer kalibrierten Gelfiltrationssäule). Nach der Verbesserung der Kristallqualität der Kristallform von TycB3-E und den ersten erfolgreichen Indizierungen wurde klar, dass in der asymmetrischen Einheitszelle ein Oligomer des rekombinanten Proteins vorliegt.

Aufgrund dieser Beobachtung und den strukturellen Erkenntnissen mit dem Protein SrfA-C stellte sich die Frage, ob der Bereich um die als COM-Domäne bezeichnete  $\alpha$ -Helix am C-Terminus vieler Epimerisierungsdomänen

1. im Falle der TycB3 Epimerisierungsdomäne für die Oligomerisierung des zunächst monodispersen Proteins in der bislang beobachteten Kristallform ursächlich ist und
2. in den vorangegangenen Kristallisationsexperimenten mit anderen, endständiger Epimerase-domänen die Entstehung von Kristallen unterbunden hat.

Die erste Frage ließ sich ohne Kenntnis der Kristallstruktur nicht beantworten. Zur Klärung der zweiten Frage sollte eine am C-Terminus verkürzte Epimerisierungsdomäne hergestellt werden. Diese COM-Deletionsmutante sollte auf ihre Kristallisierbarkeit hin untersucht werden. Zuerst wurden die C-Termini der bereits untersuchten Epimerisierungsdomänen mit Hilfe einer Sekundärstrukturvorhersage analysiert (<http://ffas.burnham.org/XtalPred-cgi/xtal.pl>).

Schließlich wurde die Epimerisierungsdomäne des Initiationsmoduls TycA ausgewählt. Mittels inverser PCR, Restriktion mit *Pst*I und anschließender Religation gelang es, den für den C-Terminus kodierenden Bereich zu entfernen und das rekombinante Protein um 35 Aminosäurereste zu verkürzen.



## 8.6 Berechnete Matthews-Koeffizienten für die Kristallform von TycB3-E

Für den Datensatz des nativen Proteins TycB3-E wurden mit dem Programm Matthews (CCP4i, 1994) die Matthews-Koeffizienten bestimmt. Wegen der Variabilität des Koeffizienten und der Größe der Einheitszelle, ist das Ergebnis nicht eindeutig. Der Bereich der möglichen Molekülanzahl reicht von 6 bis 12. Wie aus der Tabelle hervorgeht, ist es am wahrscheinlichsten, dass 9-10 Moleküle in der asymmetrischen Einheitszelle vorliegen.

Anzahl Moleküle/ asymm. Einheitszelle	Matthews-Koeff. [Å <sup>3</sup> /Da]	Solvens- gehalt [%]	Wahrscheinlichkeit (Daten bei 2.71 Å)	Wahrscheinlichkeit (alle Daten)
1	22.41	94.51	0	0
2	11.20	89.03	0	0
3	7.47	83.54	0	0
4	5.60	78.06	0	0
5	4.48	72.57	0.01	0
6	3.73	67.09	0.02	0.02
7	3.20	61.60	0.07	0.07
8	2.80	56.12	0.16	0.15
9	2.49	50.63	0.27	0.26
10	2.24	45.15	0.27	0.28
11	2.04	39.66	0.15	0.16
12	1.87	34.17	0.03	0.04

## 8.7 Versuche der Strukturlösung via *molecular replacement*

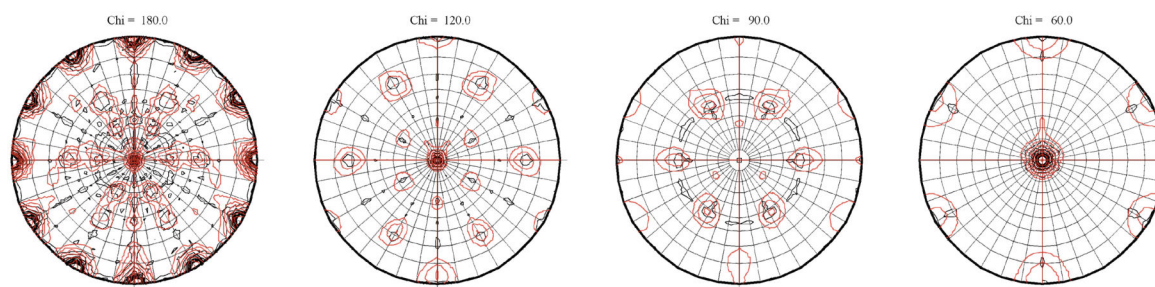
Die Methode des *molecular replacement* ist in Kapitel 3.3.7 dargestellt. Sie kann bei Vorliegen eines geeigneten Strukturmodells das Phasenproblem lösen und so eine Aufklärung der Kristallstruktur ermöglichen. Zum Gelingen der Methode ist es wichtig, dass das Modell eine große strukturelle Ähnlichkeit mit der zu lösenden Struktur aufweist, was näherungsweise anhand der Sequenzidentität beurteilt werden kann. Zu den Epimerisierungsdomänen waren zunächst keine strukturellen Informationen verfügbar. Deswegen wurde auf die Strukturen von NRPS-Kondensationsdomänen zurückgegriffen. Zwischen den E- und C-Domänen besteht, wie phylogenetische Untersuchungen [Roongsawang, 2005; Rausch, 2007] und Sekundärstrukturvorhersagen [Keating, 2002] nahelegen, eine Beziehung hinsichtlich Sequenz und Faltungstyp.

Anfangs wurden zur Generierung von Strukturmodellen die Strukturen der Kondensationsdomäne VibH (PDB: 1L5A), der oben beschriebenen Kondensationsdomäne aus TycC6 (PDB: 2JGP) und der Kondensationsdomäne aus SrfA-C (PDB: 2VSQ) [Tanovic, 2008] herangezogen. Später wurden auch die Strukturen der CoA-abhängigen Acyltransferase PapA5 (PDB: 1Q9J) [Buglino, 2004], der Acetyltransferase TRI101 (PDB: 2ZBA) [Garvey, 2008] und der TycA-E-Domäne einbezogen. Ausgehend von diesen Kristallstrukturen wurden mit dem Programm Chainsaw Strukturmodelle erzeugt. Mit diesem lassen sich einerseits basierend auf den Sequenzen (Modellprotein vs. Protein der zu lösenden Struktur) Bereiche entfernen, die voneinander abweichen, und andererseits die Seitenketten in wählbarem Maße beschneiden. Darüber hinaus wurden aus den erhaltenen Strukturmodellen in unterschiedlichem Maße solche Bereiche manuell entfernt, die gemäß Sekundärstrukturvorhersagen mit den Programmen Psipred [Jones, 1999] und RONN [Yang, 2005] als ungeordnet bezeichnet wurden. Zur Prüfung der Aussagekraft dieser beiden Programme wurden Vergleichsläufe mit der Thioesterasedomäne FenTE und der PCP-C-Bidomäne durchgeführt.

Als Kriterium für die Beurteilung der erhaltenen Lösungen wurde neben den kristallographischen Parametern auch die Selbstrotationsfunktion der Lösung herangezogen. Die Selbstrotationsfunktion der richtigen Lösung sollte, obwohl noch nicht verfeinert, dasselbe grundlegende Erscheinungsbild aufweisen wie die Selbstrotationsfunktion des Datensatzes.

Die ersten Versuche mit den Programmen Phaser und MolRep ergaben keine befriedigenden Ergebnisse. Unter Verwendung der Funktion *dyad search* im Unterprogramm *multi-copy search* des Programms MolRep wurde mit einem auf VibH basierenden Modell eine Lösung erhalten, die abgesehen von wenigen Überlappungen (*clashes*), 9 Moleküle in der asymmetrischen Einheitszelle

platzierte. Die Selbstrotationsfunktion dieser Lösung stimmte überraschend gut mit der des Datensatzes überein (Abbildung 8.1). Analoge Berechnungen, die von Strukturmodellen der beiden anderen NRPS-Kondensationsdomänen, sowie von PapA5 und TRI101 ausgingen, ergaben schlechtere Resultate, betrachtet man die deutlich abweichenden Selbstrotationsfunktionen (Daten hier nicht gezeigt).



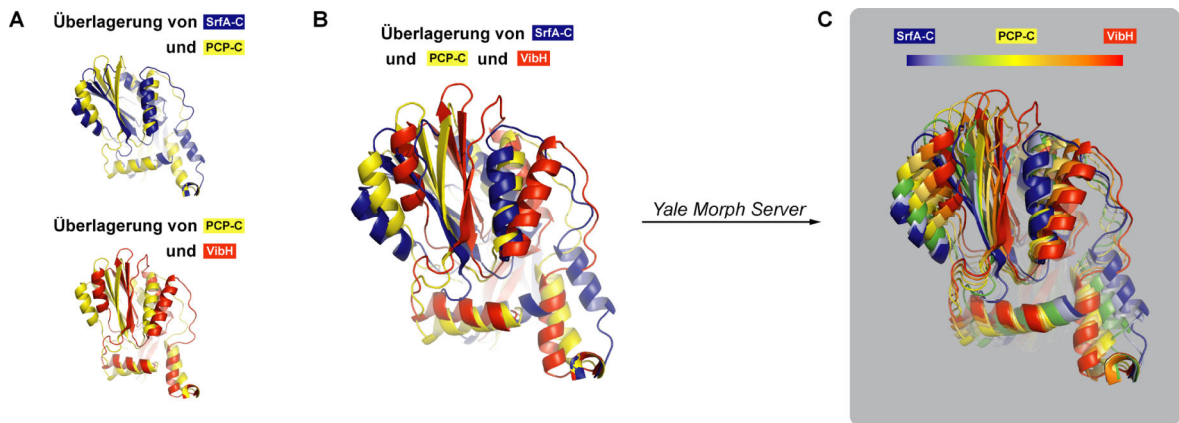
**Abbildung 8.1: Überlagerung der Patterson Selbstrotationsfunktionen des VibH-basierten Modells (rot) mit der des nativen Datensatzes von TycB3-E (schwarz)**

Die Anordnung der Strukturmodell-Moleküle gibt einige Eigenschaften der Kristallpackung richtig wieder. Allerdings ließ sich diese Lösung nicht verfeinern.

Die anfänglichen Versuche, die Struktur der PCP-C-Bidomäne mittels *molecular replacement* zu lösen, waren wohl daran gescheitert, dass die beiden Subdomänen in den beiden Proteinen unterschiedliche, relative Orientierungen einnehmen. Obwohl sich die Subdomänen in den beiden Proteinen PCP-C und VibH strukturell einander ähnlich sind, ließ sich mit dem Modell, das eine „falsche“ Anordnung zeigte, keine Lösung finden. Um einen solchen Fehler beim *molecular replacement* mit der TycB3-E-Domäne zu vermeiden, wurden weitere Strukturmodelle mit Hilfe des *Yale Morph Server* [Krebs, 2000] ([www.molmovdb.org](http://www.molmovdb.org), Januar 2008) berechnet. Dabei handelt es sich um *in silico* generierte, theoretische Zwischenzustände zwischen zwei gewählten Grenzstrukturen. So wurde zunächst die Orientierung der Subdomänen in VibH auf die PCP-C Struktur übertragen. Mit Hilfe des Servers wurde die VibH-typische Anordnung in die PCP-C-typische überführt und so 9 weitere Suchmodelle erhalten. Mit der Kondensationsdomäne der gelösten SrfA-C Struktur [Tanovic, 2008] wurde analog verfahren. Sie zeigte, verglichen mit VibH, eine entgegengesetzte Verschiebung der Subdomänen.

Die relative Verschiebung der Subdomänen zueinander ist in Abbildung 8.2A gezeigt. Dabei wurden die Proteinstrukturen auf die C-terminalen Subdomänen überlagert und jeweils nur die N-terminalen Subdomänen dargestellt. Die ursprünglichen Domänenanordnungen und eine Auswahl der künstlich erzeugten Anordnungen sind aus Abbildung 8.2B und C ersichtlich. Der Übersicht halber sind nur 7 Strukturen der N-terminalen Subdomänen dargestellt.

Die Verwendung dieser neu erzeugten Strukturmodelle in *molecular replacement*-Experimenten ergab allerdings keine weiterführenden Ergebnisse.



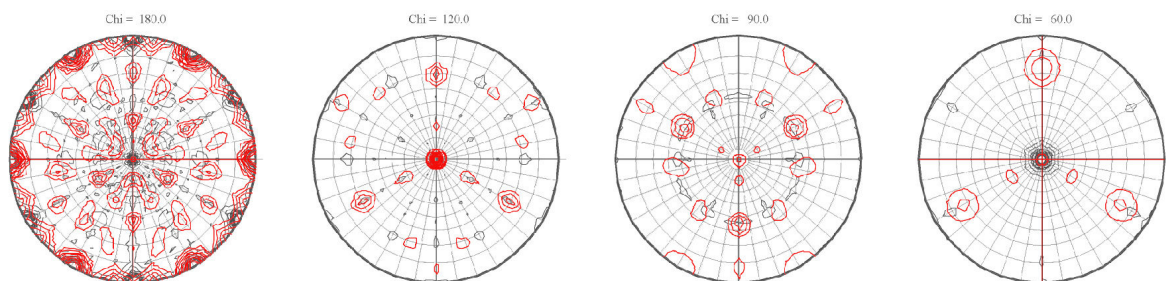
**Abbildung 8.2: Erzeugung von Struktur-Suchmodellen mit unterschiedlichen Subdomänenanordnungen**

Die Abbildung zeigt jeweils die N-terminale Subdomäne der auf die C-terminale Subdomäne überlagerten Strukturen.

(A) Unterschiedliche Orientierung der N-terminalen Subdomänen in den Strukturen von PCP-C, SrfA-C und VibH. Die Anordnung der Subdomänen in SrfA-C und VibH sind gegenüber PCP-C in entgegengesetzten Richtungen verdreht.

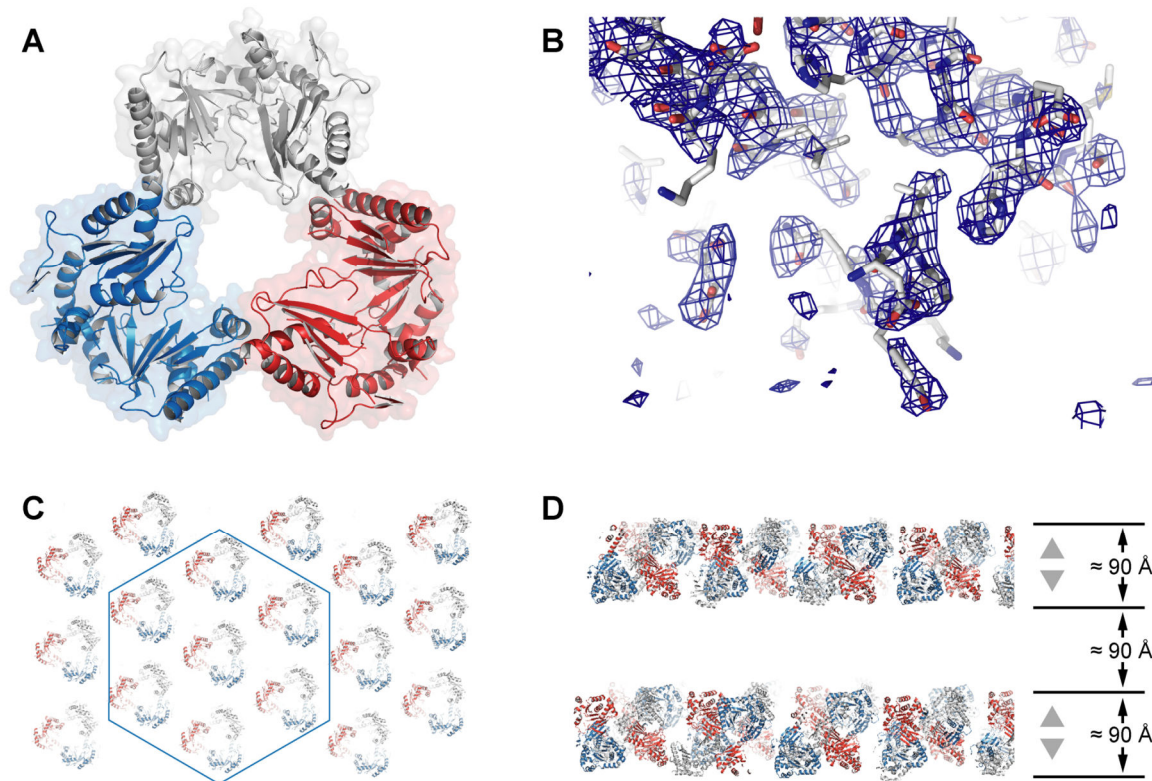
(B) Überlagerung der drei N-terminalen Subdomänen in einem Bild. (C) *In silico* generierte, theoretische Zwischenstufen von relativen Subdomänen-Anordnungen in NRPS-Kondensationsdomänen. Der Übersichtlichkeit halber sind insgesamt nur 7 Modelle abgebildet. Farbgebung der Modelle gemäß Farbverlauf-Legende.

Parallel zu diesen Arbeiten wurde die Struktur der TycA-Epimerisierungsdomäne gelöst. Aufgrund der hohen Sequenzidentität von 50 % zwischen den Epimerisierungsdomänen der Module TycA und TycB3 wurde davon ausgegangen, dass das von der TycA-E-Domäne abgeleitete Suchmodell die komplex-oligomere Struktur der TycB3-E-Domäne aufklären würde. Bedauerlicherweise führten die Versuche per *molecular replacement* bislang nicht zur Aufklärung der Struktur. Es wurde aber ein Lösungsmodell erhalten, dessen asymmetrische Einheitszelle 4-5 Moleküle enthielt. Drei dieser Moleküle zeigten eine trigonale Anordnung, wie sie anhand der Selbstrotationsfunktionen erwartet worden war. Obwohl für diese Lösung gute R-Faktoren ausgegeben wurden, gelang es nicht, die Lösung weiter zu verfeinern. Die Selbstrotationsfunktion der Lösung ist als Überlagerung mit der Referenz TycB3-E in Abbildung 8.3 dargestellt. Trotz der Unvollständigkeit der Lösung, stimmten auch für diese Lösung mehrere Signale mit der des nativen Datensatzes überein.



**Abbildung 8.3: Überlagerung der Patterson-Selbstrotationsfunktionen der mit einem TycA-E-basierten Strukturmodell erhaltenen Lösung (rot) und dem nativen Datensatz von TycB3-E (schwarz)**

Wie aus dieser Überlagerung ersichtlich, gibt die Molekülanordnung des Strukturmodells einige Eigenschaften der Kristallpackung richtig wieder.



**Abbildung 8.4: Mögliche „Teillösung“ mit trigonaler Molekülanordnung aus *molecular replacement*-Versuchen**

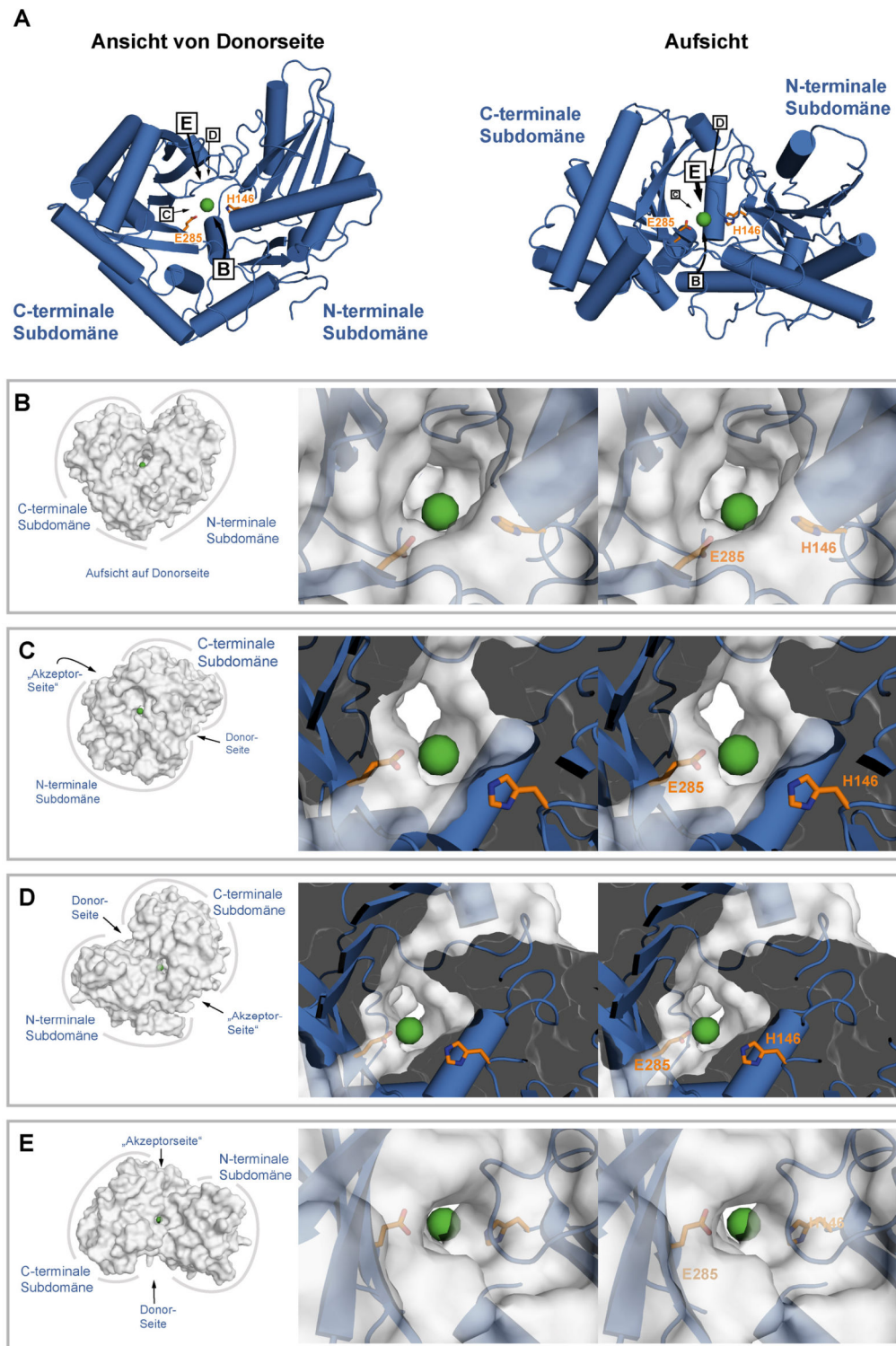
(A) Drei Moleküle des TycA-E-basierten Suchmodells in trigonaler Anordnung. (B) Der Auflösung entsprechende gute  $2F_o - F_c$ -Elektronendichte (Konturierung  $1.5 \sigma$ ) aus einer Lösung per Auto-Rickshaw. (C) Die Anordnung der Molekültrimere in der BC-Ebene stünde im Einklang mit der hexagonalen Form der Kristalle. Hier ist nur eine Schichtebene dargestellt. (D) Durch die  $C_2$ -Symmetrie entstehen gegeneinander versetzte Doppelschichten. Da das Modell aus drei Molekülen unvollständig ist, sind die Doppelschichten durch etwa  $90 \text{ \AA}$  große Leerräume voneinander getrennt. Eine solche Schichtstruktur wäre ebenfalls mit den beobachteten Fehlern in dieser Kristallform vereinbar.

Es fielen dennoch zwei Eigenschaften auf, die an dieser „Teillösung“ aus drei Proteinmolekülen bemerkenswert sind: Zum einen war der Verlauf der Hauptkette in allen drei Molekülen gut durch Elektronendichte definiert. Das ist exemplarisch in Abbildung 8.4B dargestellt. Zum anderen formten die Trimere bei Rekonstitution des Kristallgitters wabenförmige Schichten, die durch etwa ebenso große Leerräume voneinander getrennt waren. Auch wenn die Lösung nur ein unvollständiges Bild der Kristallpackung wiedergab, so ließen sich die makroskopischen Beobachtungen wie die hexagonale Form und der geschichtete Aufbau entlang der ‚hexagonalen‘ Achse gut damit vereinbaren. Unter der Annahme, dass auch im realen Kristallgitter die bindenden Wechselwirkungen zwischen den einzelnen Molekülschichten weniger stark ausgeprägt sind als innerhalb der Molekültrimere, ließe sich damit die Tendenz zur bevorzugten Ausbildung von Fehlern an den hexagonalen Kristallseiten erklären.

Auch mit der Struktur der TycA-E-Domäne, die mit der TycB3-E-Domäne eine hohe Sequenzidentität von 50 % aufweist, konnte die Molekülanordnung und die Struktur der TycB3-E-Domäne bislang nicht aufgeklärt werden.



## 8.8 Zugänge zum aktiven Zentrum der TycA-Epimerisierungsdomäne



**Abbildung 8.5: Darstellung der Öffnungen und Zugänge zum aktiven Zentrum von TycA-E**

Zur besseren Orientierung ist die Mitte des aktiven Zentrums mit einer grünen Kugel markiert. Die in den Teilen (B) bis (D) beschriebenen Zugänge sind sowohl als Totale als auch im Detail als Stereoabbildung dargestellt.

(A) Cartoonrepräsentation von TycA-E zur Übersicht der räumlichen Lage der in (B) bis (D) gezeigten Zugänge. (B) Große Öffnung auf der Donorseite. (C) Von der Donorseite betrachtet, befindet sich auf der linken Rückseite ein Kanal. (D) Auf der rechten Seite des aktiven Zentrums liegt oberhalb der  $\alpha$ -Helix  $\alpha 1$  eine Kavität sowie eine Öffnung zur Oberseite des Enzyms (im Hintergrund ist auch der in (B) gezeigte Kanal erkennbar). (E) Weitere, enge Öffnung auf der Oberseite von TycA-E.

Die in der obigen Abbildung gezeigten Zugänge zum aktiven Zentrum sind von den folgenden Bereichen umgeben:

<b><u>Öffnung zur</u></b>	<b><u>begrenzende Reste bzw. Sekundärstrukturelemente</u></b>
<b>Donorseite:</b> (Abbildung 8.5B)	Anfang <i>bridge region</i> (oben) $\alpha$ -Helix $\alpha 4$ (unten) Ende $\beta$ -Strang $\beta 9$ sowie Anfang und Ende des <i>floor loops</i> (unten)
<b>hinteren C-term. Subdomäne:</b> (Abbildung 8.5C)	Ende $\beta 12$ - $\beta 13$ Loop (oben) $\alpha$ -Helix $\alpha 10$ (unten) $\alpha$ -Helix $\alpha 1$ (unten links) $\beta$ -Stränge $\beta 9$ , $\beta 10$ und $\beta 11$ (rechts)
<b>zur hinteren Proteinoberseite:</b> (Abbildung 8.5D)	Ende <i>bridge region</i> (oben und unten) $\alpha$ -Helix $\alpha 6$ (rechts) $\beta$ -Stränge $\beta 8$ , $\beta 14$ (rechts)
<b>zur vorderen Proteinoberseite:</b> (Abbildung 8.5E)	$\beta$ -Stränge $\beta 11$ und $\beta 13$ (oben) Anfang <i>bridge region</i> (links) Reste G394 bis D396 der <i>bridge region</i> (unten)

## **Danksagung**

Die vorliegende Arbeit wurde im Zeitraum vom April 2004 bis Juli 2009 am Fachbereich Chemie der Philipps-Universität Marburg angefertigt. An dieser Stelle möchte ich mich ganz herzlich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Allen voran danke ich Herrn Prof. Dr. Lars-Oliver Essen für die Möglichkeit, dieses interessante Themengebiet zu bearbeiten, und für seine wissenschaftliche und persönliche Unterstützung während der gesamten Arbeit. Weiterhin für seine Offenheit gegenüber neuen Ideen, seine Diskussionsbereitschaft und seine guten Ratschläge, die maßgeblich zum Erfolg der angegangenen Projekte beigetragen haben.

Herrn Prof. Dr. Mohamed A. Marahiel danke ich für anregende, fachliche Diskussionen während der Promotion, hilfreiche Besprechungen bei der Erstellung von Publikationen sowie für die Übernahme des Zweitgutachtens.

Den Mitgliedern der Arbeitsgruppe Essen danke ich sowohl für die gemeinsam verbrachte Zeit im Labor, bei Seminaren und Ausflügen als auch für die angenehmen außeruniversitären Unternehmungen. Daneben danke ich auch meinen Kollegen aus den verschiedenen Kooperationsprojekten. Insbesondere bei Alan Tanovic, Thomas Knappe, Georg Schönafinger, Björn Wagner und Verena Helmetag möchte ich mich für die erfolgreiche Zusammenarbeit bedanken.

Herrn Dr. Uwe Linne danke ich sowohl für fachliche Gespräche als auch für die massenspektrometrische Untersuchung der bearbeiteten Proteine.

Außerdem bedanke ich mich bei Jürgen Behnen (Arbeitsgruppe Klebe) für die Versuche der Xenon-Bedampfung von Kristallen von TycB3-E sowie bei Dr. Andreas Heine für die Aufnahme der MAD-Datensätze, die schließlich zur Strukturlösung der Epimerisierungsdomäne führten.

Meinen ehemaligen Vertiefungspraktikanten Sevgi Ceylan, Meike Roskamp, Till Frömling, Maik Veelders, Wolfgang Große, Peter Tufar, Manuel Streib, Tobias von Mach und Xenia Vinzenz danke ich für die tatkräftige Unterstützung.

Darüber hinaus danke ich dem Graduiertenkolleg „Proteinfunktion auf atomarer Ebene“, Marburg, für die fachliche Bereicherung durch Vorträge, Seminare und Workshops, für die geselligen Studienaufenthalte im Kleinwalsertal sowie für die finanzielle Unterstützung.



Des Weiteren bedanke ich mich bei meinen Freunden Dr. Thomas Kurpiers, Dr. Georg Schönaufinger, Daniel Kreß und Marcel Sommer für viele gemeinsame Stunden während meiner Marburger Studienzeit sowie bei Daniel Sorg, Tim Hauke, Marc Stauffer und Moritz Holzmann für lustige Stunden in der Pfalz.

Meiner Freundin Nadine Bollig danke ich für all die schönen Tage, die ich bisher mit ihr verbringen durfte und ihre liebevolle Unterstützung.

Mein größter Dank gilt jedoch meinen Eltern und meiner Schwester Ines, auf deren volle Unterstützung ich mich während des gesamten Studiums und der Promotion verlassen konnte. Ihnen widme ich diese Arbeit.

## Erklärung

Ich versichere, dass ich meine Dissertation mit dem Thema:

**„Arbeiten zur strukturellen Charakterisierung von Thioesterase-,  
Kondensations- und Epimerisierungsdomänen nichtribosomaler  
Peptidsynthetasen“**

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als den von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, 01. Juli 2009

---